

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 28, 2005, 15:34:35 ; Search time 30 Seconds  
(without alignments)  
746.837 Million cell updates/sec

Title: US-10-731-554-46  
Perfect score: 1434  
Sequence: 1 MNYKRFVVGVTSTFVFLS.....ASATLNTDYFGWEIGRPFAL 271

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5 COMB.pap.\*  
2: /cgn2\_6/ptodata/1/iaa/6 COMB.pap.\*  
3: /cgn2\_6/ptodata/1/iaa/H COMB.pap.\*  
4: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pap.\*  
5: /cgn2\_6/ptodata/1/iaa/RE COMB.pap.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1434	100.0	271	2	US-09-660-587-46
2	1434	100.0	271	2	US-09-811-007A-46
3	509.5	35.5	280	2	US-09-660-587-42
4	509.5	35.5	280	2	US-09-314-701-48
5	509.5	35.5	280	2	US-09-811-007A-42
6	509.5	35.5	280	2	US-10-314-639-48
7	509.5	35.5	280	2	US-10-059-964A-48
8	507	35.4	283	2	US-09-660-587-10
9	507	35.4	283	2	US-09-261-358A-10
10	507	35.4	283	2	US-09-201-458-6
11	507	35.4	283	2	US-09-314-701-4
12	507	35.4	283	2	US-09-811-007A-10
13	507	35.4	283	2	US-10-314-639-4
14	507	35.4	283	2	US-10-059-964A-4
15	492	34.3	285	2	US-09-314-701-30
16	492	34.3	285	2	US-10-314-639-30
17	492	34.3	285	2	US-10-059-964A-30
18	489	34.1	279	2	US-09-314-701-24
19	489	34.1	279	2	US-10-314-639-24
20	489	34.1	279	2	US-10-059-964A-24
21	463	32.3	283	2	US-09-314-701-26
22	463	32.3	283	2	US-10-314-639-26
23	463	32.3	283	2	US-10-059-964A-26
24	455.5	31.8	298	2	US-09-314-701-52
25	455.5	31.8	298	2	US-10-314-639-52
26	455.5	31.8	298	2	US-10-059-964A-52
27	455	31.7	281	2	US-09-314-701-58

NO RUN

28	455	31.7	281	2	US-10-314-639-58	Sequence 58, Appl
29	455	31.7	281	2	US-10-059-964A-58	Sequence 58, Appl
30	440	30.7	291	2	US-09-314-701-18	Sequence 18, Appl
31	440	30.7	291	2	US-10-314-639-18	Sequence 18, Appl
32	440	30.7	291	2	US-10-059-964A-18	Sequence 18, Appl
33	437	30.5	293	2	US-09-660-587-40	Sequence 40, Appl
34	437	30.5	293	2	US-09-811-007A-40	Sequence 40, Appl
35	435.5	30.4	288	2	US-09-314-701-32	Sequence 32, Appl
36	435.5	30.4	288	2	US-10-314-639-32	Sequence 32, Appl
37	435.5	30.4	288	2	US-10-059-964A-32	Sequence 32, Appl
38	432	30.1	293	2	US-09-314-701-44	Sequence 44, Appl
39	432	30.1	293	2	US-10-314-639-44	Sequence 44, Appl
40	432	30.1	293	2	US-10-059-964A-44	Sequence 44, Appl
41	431.5	30.1	300	2	US-09-314-701-50	Sequence 50, Appl
42	431.5	30.1	300	2	US-10-314-639-50	Sequence 50, Appl
43	431.5	30.1	300	2	US-10-059-964A-50	Sequence 50, Appl
44	427	29.8	297	2	US-09-314-701-14	Sequence 14, Appl
45	427	29.8	297	2	US-10-314-639-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1  
US-09-660-587-46  
; Sequence 46, Application US/09660587  
; Patent No. 6392023  
; GENERAL INFORMATION:  
; APPLICANT: Walker, David H.  
; APPLICANT: McBride, Jere W.  
; APPLICANT: Yu, Xue-Jie  
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein  
; FILE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof  
; FILE REFERENCE: D6152CIP2  
; CURRENT APPLICATION NUMBER: US/09/660,587  
; CURRENT FILING DATE: 2000-09-12  
; PRIOR APPLICATION NUMBER: 09/261,358  
; PRIOR FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 46  
; SEQ ID NO 46  
; LENGTH: 271  
; TYPE: PRT  
; ORGANISM: Ehrlichia canis  
; FEATURE:  
; OTHER INFORMATION: amino acid sequence of E. canis p28-9 protein  
US-09-660-587-46

Query Match	100.0%	Score 1434;	DB 2;	Length 271;
Best Local Similarity	100.0%;	Pred. No. 2.3e-147;	Mismatches 0;	Indels 0; Gaps 0;
Matches 271;	Conservative 0;			
Qy	1	MNYKRFVVGVTSTFVFLSDGAFSDANFSEGRRLYIGSQYKVGIPNFNFSAEETIPG	60	
Db	1	MNYKRFVVGVTSTFVFLSDGAFSDANFSEGRRLYIGSQYKVGIPNFNFSAEETIPG	60	
Qy	61	ITKKIFALGDKSEINTHSNFTSYDPTASSFAGSGIIGYVNDFRVEFGSYENPEP	120	
Db	61	ITKKIFALGDKSEINTHSNFTSYDPTASSFAGSGIIGYVNDFRVEFGSYENPEP	120	
Qy	121	EROWYPENSOSYKFFALSRLNATNSDKFIVLENNGVVDKSLNNVVCYDIASGSIPLAPYM	180	
Db	121	EROWYPENSOSYKFFALSRLNATNSDKFIVLENNGVVDKSLNNVVCYDIASGSIPLAPYM	180	
Qy	181	CAGVGADYIKFLGISTLPKFSYQVKFGVYPLNVNTMLFGGGYHKVVGDRHERVEIAYHP	240	
Db	181	CAGVGADYIKFLGISTLPKFSYQVKFGVYPLNVNTMLFGGGYHKVVGDRHERVEIAYHP	240	
Qy	241	TALSDVPRTTTSASATLNTDYFGWEIGRPFAL	271	
Db	241	TALSDVPRTTTSASATLNTDYFGWEIGRPFAL	271	

RESULT 2

```
US-09-811-007A-46
; Sequence 46, Application US/09811007A
; Patent No. 6660269
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
; FILE REFERENCE: D6152CIP2
; CURRENT APPLICATION NUMBER: US/09/811,007A
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 09/660,587
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 46
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: amino acid sequence of E. canis p28-9 protein
US-09-811-007A-46

Query Match      100.0%; Score 1434; DB 2; Length 271;
Best Local Similarity 100.0%; Pred. No. 2.3e-147;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNYKRFVVGVTLSFTVFFFLSDGAFSDANFSEGRRLYIGSQYKVGIPNFSNFSAEETIPG 60
Db 1 MNYKRFVVGVTLSFTVFFFLSDGAFSDANFSEGRRLYIGSQYKVGIPNFSNFSAEETIPG 60

QY 61 ITKKIFALGLDKSEINTHSNFTSYDPTVASSFAGSGIIGYVNDPFRVEFEGSYENFEP 120
Db 61 ITKKIFALGLDKSEINTHSNFTSYDPTVASSFAGSGIIGYVNDPFRVEFEGSYENFEP 120

QY 121 ERQWYPENSQSYKFFALSRNATNSDNKFIVLNNGVVVDSKSLNVNVCYDIASGSIPLAPYM 180
Db 121 ERQWYPENSQSYKFFALSRNATNSDNKFIVLNNGVVVDSKSLNVNVCYDIASGSIPLAPYM 180

QY 181 CAGVGADYIKFLGISLPKFSYQVKGFGVNYPLNVNTMLFGGGYHKVVGDRHVERVEIAYHP 240
Db 181 CAGVGADYIKFLGISLPKFSYQVKGFGVNYPLNVNTMLFGGGYHKVVGDRHVERVEIAYHP 240

QY 241 TALSDVPRTTSSASATLNTDYFGWEIGRFPAL 271
Db 241 TALSDVPRTTSSASATLNTDYFGWEIGRFPAL 271

RESULT 3
US-09-660-587-42
; Sequence 42, Application US/09660587
; Patent No. 6392023
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
; FILE REFERENCE: D6152CIP2
; CURRENT APPLICATION NUMBER: US/09/660,587
; CURRENT FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 09/261,358
; PRIOR FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 42
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: amino acid sequence of E. canis p28-2 protein
US-09-660-587-42

Query Match      35.5%; Score 509.5; DB 2; Length 280;
Best Local Similarity 40.5%; Pred. No. 7.5e-47;
Matches 115; Conservative 50; Mismatches 98; Indels 21; Gaps 10;

QY 1 MNYKRFVVGVTLSFTVFFFLSDGAFSD---ANFSEGRRLYIGSQYKVGIPNFSNFSAEET 57
Db 1 MNYKKILVRSALISLSILPYQSFPADPVGSRDNDNKEGFYISAKYNPSISHFRKFSAEET 60

QY 58 -IPG---ITKKIFALGLDKSEINTHSNFT---SYDPTVASSFAGSGIIGYVNDPFR 108
Db 61 PINGTNSLTAKVFGGLKKD-GDITKKDDFTRVAPGIDFQNNLIS---GFSGSIGYSMDGPR 116

QY 109 VEFEGSYENFEPER--QWYPENSQSYKFFALSRNATNSDNKFIVLNNGVVVDSKSLNVNVC 166
Db 117 IELEAAYQQFNPKNNTDNDTNGEYKHFALSRKDAMEDQQVYVVKNDGITTMSLMVNTC 176

QY 167 YDIASGSIPLAPYMCAGVADYIK-FLGISLPKFSYQVKGFGVNYPLNVNTMLFGGGYHK 225
Db 177 YDITAEGVSVFVPYACAGIGADLITIFKDLNL-KFAYQKGIGISYPTTPVPSAFIGGYHG 235

QY 226 VVGDRHVERVEIAYHPTALSDVPRTTSSASATLNTDYFGWEIGRPF 269
Db 236 VIGNKFEKIPV-ITPVLNDAPQTTASVTLTDVGYPFGGEIGMRF 278

RESULT 4
US-09-314-701-48
; Sequence 48, Application US/09314701
; Patent No. 6544517
; GENERAL INFORMATION:
; APPLICANT: Rikihisa, Yasuko
; APPLICANT: Ohasi, No. 6544517io
; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
; TITLE OF INVENTION: Chaffeensis
; FILE REFERENCE: 22727/04021
; CURRENT APPLICATION NUMBER: US/09/314,701
; CURRENT FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Ehrlichia canis
; ORGANISM: Ehrlichia canis
US-09-314-701-48

Query Match      35.5%; Score 509.5; DB 2; Length 280;
Best Local Similarity 40.5%; Pred. No. 7.5e-47;
Matches 115; Conservative 50; Mismatches 98; Indels 21; Gaps 10;

QY 1 MNYKRFVVGVTLSFTVFFFLSDGAFSD---ANFSEGRRLYIGSQYKVGIPNFSNFSAEET 57
Db 1 MNYKKILVRSALISLSILPYQSFPADPVGSRDNDNKEGFYISAKYNPSISHFRKFSAEET 60

QY 58 -IPG---ITKKIFALGLDKSEINTHSNFT---SYDPTVASSFAGSGIIGYVNDPFR 108
Db 61 PINGTNSLTAKVFGGLKKD-GDITKKDDFTRVAPGIDFQNNLIS---GFSGSIGYSMDGPR 116

QY 109 VEFEGSYENFEPER--QWYPENSQSYKFFALSRNATNSDNKFIVLNNGVVVDSKSLNVNVC 166
Db 117 IELEAAYQQFNPKNNTDNDTNGEYKHFALSRKDAMEDQQVYVVKNDGITTMSLMVNTC 176

QY 167 YDIASGSIPLAPYMCAGVADYIK-FLGISLPKFSYQVKGFGVNYPLNVNTMLFGGGYHK 225
Db 177 YDITAEGVSVFVPYACAGIGADLITIFKDLNL-KFAYQKGIGISYPTTPVPSAFIGGYHG 235

QY 226 VVGDRHVERVEIAYHPTALSDVPRTTSSASATLNTDYFGWEIGRPF 269
Db 236 VIGNKFEKIPV-ITPVLNDAPQTTASVTLTDVGYPFGGEIGMRF 278

RESULT 5
US-09-811-007A-42
; Sequence 42, Application US/09811007A
; Patent No. 6660269
```



; APPLICANT: Walker, David H.  
; APPLICANT: McBride, Jere W.  
; APPLICANT: Yu, Xue-Jie  
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein  
; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof  
; FILE REFERENCE: D6152C1P2  
; CURRENT APPLICATION NUMBER: US/09/660,587  
; PRIOR FILING DATE: 2000-09-12  
; PRIOR APPLICATION NUMBER: 09/261,358  
; PRIOR FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 46  
; SEQ ID NO 10  
; LENGTH: 283  
; TYPE: PRT  
; ORGANISM: Ehrlichia chaffeensis  
; FEATURE:  
; OTHER INFORMATION: amino acid sequence of E. chaffeensis OMP-1B  
US-09-660-587-10

Query Match 35.4%; Score 507; DB 2; Length 283;  
Best Local Similarity 40.3%; Pred. No. 1.4e-46;  
Matches 116; Conservative 46; Mismatches 100; Indels 26; Gaps 11;

QY 1 MNYKRFVVGVTLSFVFFLSDGAF-----SDANFSEGRRLYIGSQYKVGIPNFSNFA 54  
Db 1 MNYKKIFVSSALISLSILPYQSFADPVTSTNDTGINDSREGFYISVKYNPSISHPRKPSA 60  
QY 55 EET-IPG---ITKKIFALGLDKSEINHTSNFTSRSDPTYA---SSPAGSGIIGYVNDP 107  
Db 61 EAPINGNTSITKKVFLGKKD-GDIAQSANFNRT-DPALEFQNNLISGFSGISYAMDGP 118  
QY 108 RVFEFGSYENFPERQWYPENSOS-----YKFPALSRNATNSDNKFIVLNNGVVDKSLN 162  
Db 119 RIELEAAVQKFDKXN---PDNDTNSGDYKYFGLSREDAIADKKVVLKNEGIFMMSLM 175  
QY 163 VNVCDYIAGSIPAPYMCAGVADYIK-FLGISLPKFSYQVKGVNYPNVTMLFGGG 221  
Db 176 VNTCYDITAGVFPFIPYACAGVGADLINVKDFNL-KFSYQKGIGISYIPITPEVSAFIGG 234  
QY 222 YHKVVGDRHERVEIAYHPTALSDVPRTTTSASATLNTDYFGWEIGRPF 269  
Db 235 YYHGVIGNNFNPKIPV-ITPVVLEGAPQTTTSALVTIDTGYFGGEGVGRF 281

RESULT 9  
US-09-261-358A-10  
; Sequence 10, Application US/09261358A  
; Patent No. 6403780  
; GENERAL INFORMATION:  
; APPLICANT: Walker, David H.  
; APPLICANT: McBride, Jere W.  
; APPLICANT: Yu, Xue-Jie  
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein  
; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof  
; FILE REFERENCE: D6152C1P  
; CURRENT APPLICATION NUMBER: US/09/261,358A  
; CURRENT FILING DATE: 1999-03-03  
; PRIOR APPLICATION NUMBER: 09/201,458  
; PRIOR FILING DATE: 1998-11-30  
; NUMBER OF SEQ ID NOS: 33  
; SEQ ID NO 10  
; LENGTH: 283  
; TYPE: PRT  
; ORGANISM: Ehrlichia chaffeensis  
; FEATURE:  
; OTHER INFORMATION: amino acid sequence of E. chaffeensis OMP-1B  
US-09-261-358A-10

Query Match 35.4%; Score 507; DB 2; Length 283;  
Best Local Similarity 40.3%; Pred. No. 1.4e-46;  
Matches 116; Conservative 46; Mismatches 100; Indels 26; Gaps 11;

QY 1 MNYKRFVVGVTLSFVFFLSDGAF-----SDANFSEGRRLYIGSQYKVGIPNFSNFA 54

Db 1 MNYKKIFVSSALISLSILPYQSFADPVTSTNDTGINDSREGFYISVKYNPSISHPRKPSA 60  
QY 55 EET-IPG---ITKKIFALGLDKSEINHTSNFTSRSDPTYA---SSPAGSGIIGYVNDP 107  
Db 61 EAPINGNTSITKKVFLGKKD-GDIAQSANFNRT-DPALEFQNNLISGFSGISYAMDGP 118  
QY 108 RVFEFGSYENFPERQWYPENSOS-----YKFPALSRNATNSDNKFIVLNNGVVDKSLN 162  
Db 119 RIELEAAVQKFDKXN---PDNDTNSGDYKYFGLSREDAIADKKVVLKNEGIFMMSLM 175  
QY 163 VNVCDYIAGSIPAPYMCAGVADYIK-FLGISLPKFSYQVKGVNYPNVTMLFGGG 221  
Db 176 VNTCYDITAGVFPFIPYACAGVGADLINVKDFNL-KFSYQKGIGISYIPITPEVSAFIGG 234  
QY 222 YHKVVGDRHERVEIAYHPTALSDVPRTTTSASATLNTDYFGWEIGRPF 269  
Db 235 YYHGVIGNNFNPKIPV-ITPVVLEGAPQTTTSALVTIDTGYFGGEGVGRF 281

RESULT 10  
US-09-201-458-6  
; Sequence 6, Application US/09201458A  
; Patent No. 6458942  
; GENERAL INFORMATION:  
; APPLICANT: Walker, David H.  
; APPLICANT: McBride, Jere W.  
; APPLICANT: Yu, Xue-Jie  
; TITLE OF INVENTION: 28-kDa Immureactive Protein Gene of Ehrlichia  
; TITLE OF INVENTION: canis and Uses Thereof  
; FILE REFERENCE: D6152  
; CURRENT APPLICATION NUMBER: US/09/201,458A  
; CURRENT FILING DATE: 1998-11-30  
; NUMBER OF SEQ ID NOS: 21  
; SEQ ID NO 6  
; LENGTH: 283  
; TYPE: PRT  
; ORGANISM: Ehrlichia chaffeensis  
; FEATURE:  
; OTHER INFORMATION: amino acid sequence of E. chaffeensis OMP-1B  
US-09-201-458-6

Query Match 35.4%; Score 507; DB 2; Length 283;  
Best Local Similarity 40.3%; Pred. No. 1.4e-46;  
Matches 116; Conservative 46; Mismatches 100; Indels 26; Gaps 11;

QY 1 MNYKRFVVGVTLSFVFFLSDGAF-----SDANFSEGRRLYIGSQYKVGIPNFSNFA 54  
Db 1 MNYKKIFVSSALISLSILPYQSFADPVTSTNDTGINDSREGFYISVKYNPSISHPRKPSA 60  
QY 55 EET-IPG---ITKKIFALGLDKSEINHTSNFTSRSDPTYA---SSPAGSGIIGYVNDP 107  
Db 61 EAPINGNTSITKKVFLGKKD-GDIAQSANFNRT-DPALEFQNNLISGFSGISYAMDGP 118  
QY 108 RVFEFGSYENFPERQWYPENSOS-----YKFPALSRNATNSDNKFIVLNNGVVDKSLN 162  
Db 119 RIELEAAVQKFDKXN---PDNDTNSGDYKYFGLSREDAIADKKVVLKNEGIFMMSLM 175  
QY 163 VNVCDYIAGSIPAPYMCAGVADYIK-FLGISLPKFSYQVKGVNYPNVTMLFGGG 221  
Db 176 VNTCYDITAGVFPFIPYACAGVGADLINVKDFNL-KFSYQKGIGISYIPITPEVSAFIGG 234  
QY 222 YHKVVGDRHERVEIAYHPTALSDVPRTTTSASATLNTDYFGWEIGRPF 269  
Db 235 YYHGVIGNNFNPKIPV-ITPVVLEGAPQTTTSALVTIDTGYFGGEGVGRF 281

RESULT 11  
US-09-314-701-4  
; Sequence 4, Application US/09314701  
; Patent No. 6544517  
; GENERAL INFORMATION:  
; APPLICANT: Rikihisa, Yasuko





```
; CURRENT FILING DATE: 2002-01-28
; PRIOR APPLICATION NUMBER: 09/314,701
; PRIOR FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: 60/100,843
; PRIOR FILING DATE: 1998-09-18
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: Patent In Ver. 3.2
; SEQ ID NO 4
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Ehrlichia chaffeensis
US-10-059-964A-4

Query Match          35.4%; Score 507; DB 2; Length 283;
Best Local Similarity 40.3%; Pred. No. 1.4e-46;
Matches 116; Conservative 46; Mismatches 100; Indels 26; Gaps 11;

Qy 1 MNYKRFVGVGTLSTFVFFLSDGAF-----SDANFSEGRGLYIGSQYKVGIPNFSNFA 54
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MNYKIFVSSALISLMSILPYQSFADPVTSTNDTGINDSREGFYISVKYNPSISHFRKFA 60

Qy 55 EET-IPG---ITKKIFALGLDKSEINTHSNFTRSYDPTYA---SSPAGSGLIIGYYVND 107
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 EEAPINGNTSITKKVFLGKXD-GDIAQSANFNRT-DPALEFQNNLISGFSGSIGYAMDGP 118

Qy 108 RVFEFGSYENFERQWYPENSOS-----YKFFALSERNATNSDNKFTVLENNGVVDKSLN 162
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 RIBLEAAQYQFDKDN---PDNNDTNSGDYKYFGLGREDAIAOKKYVVLKNEGITPMWSLM 175

Qy 163 VNVCYDIAGSIPLAPYMCAGVGADYIK-FLGISLPKFSYQVKFGVNYPLNVNNTMLFGGG 221
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 176 VNTCYDITASGVFPPIPVACAGVGADLINVPKFNL-KFSYQKGIGISYIPITPEVSFIGG 234

Qy 222 YVHKVVDGRHERVEIAVHPTALSDVPRTTTSASATLNTDYPGWGIGRFRF 269
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 235 YVHGVIQNNFNKIPV-ITPVVLEGAPQTTTSALVTIDTGYFEGSGVGVRF 281
```

```
RESULT 15
US-09-314-701-30
; Sequence 30, Application US/09314701
; Patent No. 6544517
; GENERAL INFORMATION:
; APPLICANT: Rikihisa, Yasuko
; APPLICANT: Ohasi, No. 6544517io
; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
; FILE REFERENCE: Chaffeensis
; CURRENT APPLICATION NUMBER: US/09/314,701
; CURRENT FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 30
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Ehrlichia chaffeensis
US-09-314-701-30

Query Match          34.3%; Score 492; DB 2; Length 285;
Best Local Similarity 38.4%; Pred. No. 6.1e-45;
Matches 111; Conservative 56; Mismatches 96; Indels 26; Gaps 10;

Qy 1 MNYKR--FVGVGTLSTFVFFLSDGAFSD--ANFSEGRGLYIGSQYKVGIPNFSNFAEE 56
Db 1 MNNRKFFIIIGASLLASLFTSASTGNSGVNSNHTYKPRLYISGQYRPGVSHFSKFSVKE 60

Qy 57 T-----IPGITKKIFALGLDKSEINTHSNFTRSYDPTYASSPAGSGLIIGY-YVNDFRV 109
Db 61 TNYNTQLVGLKKDLSVIG--NSNITYTNFNFFYIAEQDNAISFGAIGLYISENFRI 118

Qy 110 EFGSGYENFERQWYPENS--QSYKFFALSR-----NATNSDN--KFIVLENNGVVDKXS 160
Db 119 EVEASYEEFDVKN---PEGSATDAYRYFALARAMDGTNKSPPDTRKFTVMRNDGLSISS 175
```

```
Qy 161 LNVNVCYDIAGSIPLAPYMCAGVGADYIKFLGISLPKFSYQVKFGVNYPLNVNNTMLFGG 220
Db 176 VNINGCYNFTLDDIPVVPYVCAGIGGDFIEFFNDLHVKFRHQQKVGISYSISPEVSLFLN 235
Qy 221 GYYHKVVDGRHERVEIAVHPTALSDVPRTTTSASATLNTDYPGWGIGRFRF 269
Db 236 GYYHKVTGNRFKNLHV-QHVSDLSADAPKFTSAVATLNVGYFGEIGVRF 283
```

Search completed: December 28, 2005, 15:43:04  
Job time : 31 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 28, 2005, 15:02:01 ; Search time 111 Seconds  
(without alignments)

1722.506 Million cell updates/sec

Title: US-10-731-554-46

Perfect score: 1434

Sequence: 1 MNYKRFVGVGTLSTFVPLS.....ASATLNTDYFGWEIGFRPAL 271

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1434	100.0	271	2	Q9F471_EHRCA	Q9f471 ehrlichia c
2	1424	99.3	275	2	Q9ADV2_EHRCA	Q9adv2 ehrlichia c
3	1194	83.3	271	2	O69197_EHRCH	O69197 ehrlichia c
4	903	63.0	283	2	O5FGQ6_EHRRG	O5fgq6 ehrlichia r
5	903	63.0	283	2	Q4LOB8_EHRRU	Q4lob8 ehrlichia r
6	892	62.2	285	2	Q6VCX5_EHRRU	Q6vcx5 ehrlichia r
7	892	62.2	285	2	Q5HA08_EHRRW	Q5ha08 ehrlichia r
8	856	59.7	271	2	Q4W4W3_EHRRU	Q4w4w3 ehrlichia r
9	525.5	36.6	282	2	O5W5Y1_EHRRU	O5w5y1 ehrlichia r
10	522.5	36.4	282	2	Q5HA10_EHRRW	Q5ha10 ehrlichia r
11	522.5	36.4	282	2	Q9W4W1_EHRRU	Q9w4w1 ehrlichia r
12	517.5	36.1	282	2	Q9R443_EHRRU	Q9r443 ehrlichia r
13	515	35.9	283	2	Q8G812_EHRCH	Q8g812 ehrlichia c
14	514.5	35.9	282	2	Q4LOC0_EHRRU	Q4loc0 ehrlichia r
15	514.5	35.9	304	2	O5FGL4_EHRRG	O5fgl4 ehrlichia r
16	514.5	35.9	304	2	O5FCD9_EHRRW	O5fcd9 ehrlichia r
17	514	35.8	283	2	Q8G8D6_EHRCH	Q8g8d6 ehrlichia c
18	509.5	35.5	280	2	Q9F476_EHRCA	Q9f476 ehrlichia c
19	507	35.4	283	2	O8GGU4_EHRCH	O8ggu4 ehrlichia c
20	507	35.4	283	2	O85358_EHRCH	O85358 ehrlichia c
21	504.5	35.2	280	2	Q84HU1_EHRCA	Q84hu1 ehrlichia c
22	494	34.4	285	2	O9L6V5_EHRCH	O9l6v5 ehrlichia c
23	489	34.1	279	2	Q9ACJ2_EHRCH	Q9acj2 ehrlichia c
24	482	33.6	279	2	Q9L6Y9_EHRCH	Q9l6y9 ehrlichia c
25	469	32.7	282	2	Q4W4X0_EHRRU	Q4w4x0 ehrlichia r
26	467	32.6	282	2	O5FGM2_EHRRG	O5fgm2 ehrlichia r
27	467	32.6	282	2	Q4LOC7_EHRRU	Q4loc7 ehrlichia r
28	464	32.4	282	2	O5HA16_EHRRW	O5ha16 ehrlichia r
29	463	32.3	283	2	Q9L6Y7_EHRCH	Q9l6y7 ehrlichia c
30	460	32.1	281	2	Q9ADV7_EHRCA	Q9adv7 ehrlichia c
31	460	32.1	282	2	Q6VCY4_EHRRU	Q6vcy4 ehrlichia r

32	455.5	31.8	298	2	Q9L6V3_EHRCH	Q9l6v3 ehrlichia c
33	447.5	31.2	300	2	Q9L6V2_EHRCH	Q9l6v2 ehrlichia c
34	446	31.1	289	2	Q9ADV9_EHRCA	Q9adv9 ehrlichia c
35	442.5	30.9	294	2	Q9ADV4_EHRCA	Q9adv4 ehrlichia c
36	442	30.8	278	2	Q9ADW1_EHRCA	Q9adw1 ehrlichia c
37	440	30.7	291	2	Q9ACJ0_EHRCH	Q9acj0 ehrlichia c
38	437.5	30.5	290	2	O5FGL6_EHRRG	O5fgl6 ehrlichia r
39	437.5	30.5	290	2	O4LOC1_EHRRU	O4loc1 ehrlichia r
40	437	30.5	293	2	Q9F477_EHRCA	Q9f477 ehrlichia c
41	436	30.4	275	2	Q93DD4_EHRCH	Q93dd4 ehrlichia c
42	435.5	30.4	288	2	Q9ZGJ2_EHRCA	Q9zgj2 ehrlichia c
43	428	29.8	283	2	O4LOC6_EHRRU	O4loc6 ehrlichia r
44	428	29.8	283	2	O6VCY3_EHRRU	O6vcy3 ehrlichia r
45	428	29.8	283	2	Q5HA15_EHRRW	Q5ha15 ehrlichia r

ALIGNMENTS

RESULT 1  
Q9F471\_EHRCA  
ID Q9F471\_EHRCA PRELIMINARY; PRT; 271 AA.  
AC Q9F471;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE P28-9.  
GN Name=p28-9;  
OS Ehrlichia canis.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;  
OC Anaplasmataceae; Ehrlichia.  
OX NCBI\_TaxID=944;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=Juke;  
RX MEDLINE=99242757; PubMed=10225842;  
RA McBride J.W., Yu X.J., Walker D.H.;  
RT "Molecular cloning of the gene for a conserved major immunoreactive 28-kilodalton protein of Ehrlichia canis: a potential serodiagnostic antigen."  
RT Clin. Diagn. Immunol. 6:392-399(1999).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=Juke;  
RX MEDLINE=20432107; PubMed=10974556; DOI=10.1016/S0378-1119(00)00256-0;  
RA McBride J.W., Yu X.J., Walker D.H.;  
RT "A conserved, transcriptionally active p28 multigene locus of Ehrlichia canis".  
RT Gene 254:245-252(2000).  
DR EMBL; AF082744; AAG14363.1; -; Genomic DNA.  
DR InterPro; IPR002566; Surface\_Ag\_msp4.  
DR Pfam; PF01617; Surface\_Ag\_2; 1.  
SQ SEQUENCE 271 AA; 30374 MW; 8C8F22BB1D05B3EE CRC64;

Query Match 100.0%; Score 1434; DB 2; Length 271;  
Best Local Similarity 100.0%; Pred. No. 1.2e-109;  
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNYKRFVGVGTLSTFVFLSDGAFSDANFSEGRGLYIGSQYKVGIPNFSNFSAEETIPG 60  
DB 1 MNYKRFVGVGTLSTFVFLSDGAFSDANFSEGRGLYIGSQYKVGIPNFSNFSAEETIPG 60  
QY 61 ITTKIFALGLDKSEINHSNTRSDPTVYASSFAGSGIIGYVNDFRVFEFGSVENFEP 120  
DB 61 ITTKIFALGLDKSEINHSNTRSDPTVYASSFAGSGIIGYVNDFRVFEFGSVENFEP 120  
QY 121 ERQWTPENSQSYKFFALSRNATNSDNKFTVLENNGVDKSLNVNVCYDIASGSIPLAPYM 180  
DB 121 ERQWTPENSQSYKFFALSRNATNSDNKFTVLENNGVDKSLNVNVCYDIASGSIPLAPYM 180  
QY 181 CAGVGADYIKFLGIGISLPKFSQVKGVNYPLNVNTMLFGGYHYHKVGVGRHVERVEYHP 240  
DB 181 CAGVGADYIKFLGIGISLPKFSQVKGVNYPLNVNTMLFGGYHYHKVGVGRHVERVEYHP 240

```

QY 241 TALSDVPRTTSASATLNTDYFGWIGRFRFAL 271
DB 241 TALSDVPRTTSASATLNTDYFGWIGRFRFAL 271

RESULT 2
Q9ADV2_EHRCA PRELIMINARY; PRT; 275 AA.
AC Q9ADV2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Major outer membrane protein P30-20.
GN Name=p30-20;
OS Ehrlichia canis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Ehrlichia.
OX NCBI_TaxID=944;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Oklahoma;
RX MEDLINE=98371112; PubMed=9705412;
RA Ohashi N., Unver A., Zhi N., Rikihisa Y.;
RT "Cloning and characterization of multigenes encoding the
RT immunodominant 30-kilodalton major outer membrane proteins of
RT Ehrlichia canis and application of the recombinant protein for
RT serodiagnosis.";
RL J. Clin. Microbiol. 36:2671-2680 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Oklahoma;
RX MEDLINE=21153566; PubMed=11254561;
RA Ohashi N., Rikihisa Y., Unver A.;
RT "Analysis of transcriptionally active gene clusters of major outer
RT membrane protein multigene family in Ehrlichia canis and E.
RT chaffeensis.";
RL Infect. Immun. 69:2083-2091 (2001).
DR EMBL; AF078553; AAK28700.1; -; Genomic_DNA.
DR InterPro; IPR002566; Surface_Ag_msp4.
DR Pfam; PF01617; Surface_Ag_2; 1.
SQ SEQUENCE 275 AA; 30843 MW; 3D8FD8FE507CE51 CRC64;

Query Match 99.3%; Score 1424; DB 2; Length 275;
Best Local Similarity 99.3%; Pred. No. 7.8e-109;
Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNYKRFVVGVTLSFTVFPLSDGAFSDANFSEGRGLYIGSQYKVGIPNFSNFSAEETIPG 60
DB 5 MNYKRFVVGVTLSFTVFPLSDGAFSDANFSEGRGLYIGSQYKVGIPNFSNFSAEETIPG 64
QY 61 ITKKIFALGLDKSEINTHSNFTSRSDPTYASSPAGFSGIIGYVNDPRVFEFGSYENFEP 120
DB 65 ITKKIFALGLDKSEINTHSNFTSRSDPTYASSPAGFSGIIGYVNDPRVFEFGSYENFEP 124
QY 121 ERQWYPENSQSYKFFALSRLNATSDNKFIVLENNGVVDKSLNNVCYDIASGSIPLAPYM 180
DB 125 ERQWYPENSQSYKFFALSRLNATSDNKFIVLENNGVVDKSLNNVCYDIASGSIPLAPYM 184
QY 181 CAGVGADYIKFLGISLKPFSYQVKGNYPLNVTMLFGGGYHKVGDHRRVERVEIAYHP 240
DB 185 CAGVGADYIKFLGISLKPFSYQVKGNYPLNVTMLFGGGYHKVGDHRRVERVEIAYHP 244
QY 241 TALSDVPRTTSASATLNTDYFGWIGRFRFAL 271
DB 245 TALSDVPRTTSASATLNTDYFGWIGRFRFAL 275

RESULT 3
O69197_EHRCH PRELIMINARY; PRT; 271 AA.
AC O69197;

QY 241 TALSDVPRTTSASATLNTDYFGWIGRFRFAL 271
DB 241 TALSDVPRTTSASATLNTDYFGWIGRFRFAL 271

RESULT 4
Q5FGQ6_EHRRG PRELIMINARY; PRT; 283 AA.
AC Q5FGQ6;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)

```

```

DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE P28-20 precursor (Major outer membrane protein p28-1).
GN Name=p28-20; Synonyms=p28-1;
OS Ehrlichia chaffeensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Ehrlichia.
OX NCBI_TaxID=945;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Arkansas;
RX MEDLINE=99175287; PubMed=10074538;
RA Yu X.J., McBride J.W., Walker D.H.;
RT "Genetic diversity of the 28-kilodalton outer membrane protein gene in
RT human isolates of Ehrlichia chaffeensis.";
RL J. Clin. Microbiol. 37:1137-1143 (1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Arkansas;
RX MEDLINE=21153566; PubMed=11254561;
RA Ohashi N., Rikihisa Y., Unver A.;
RT "Analysis of transcriptionally active gene clusters of major outer
RT membrane protein multigene family in Ehrlichia canis and E.
RT chaffeensis.";
RL Infect. Immun. 69:2083-2091 (2001).
DR EMBL; AF068234; AAC19134.2; -; Genomic_DNA.
DR EMBL; U72291; AAK28674.1; -; Genomic_DNA.
DR InterPro; IPR002566; Surface_Ag_msp4.
DR Pfam; PF01617; Surface_Ag_2; 1.
KW Signal.
FT SIGNAL 1 25 Potential.
FT CHAIN 26 271 P28-20.
SQ SEQUENCE 271 AA; 30050 MW; 84EDF665DF3DB61C CRC64;

Query Match 83.3%; Score 1194; DB 2; Length 271;
Best Local Similarity 78.6%; Pred. No. 6.2e-90;
Matches 213; Conservative 33; Mismatches 25; Indels 0; Gaps 0;

QY 1 MNYKRFVVGVTLSFTVFPLSDGAFSDANFSEGRGLYIGSQYKVGIPNFSNFSAEETIPG 60
DB 1 MNYKRFVVGVTLSFTVFPLSDGAFSDANFSEGRGLYIGSQYKVGIPNFSNFSAEETIPG 60
QY 61 ITKKIFALGLDKSEINTHSNFTSRSDPTYASSPAGFSGIIGYVNDPRVFEFGSYENFEP 120
DB 61 ITKKIFALGLDKSEINTHSNFTSRSDPTYASSPAGFSGIIGYVNDPRVFEFGSYENFEP 120
QY 121 ERQWYPENSQSYKFFALSRLNATSDNKFIVLENNGVVDKSLNNVCYDIASGSIPLAPYM 180
DB 121 ERQWYPENSQSYKFFALSRLNATSDNKFIVLENNGVVDKSLNNVCYDIASGSIPLAPYM 180
QY 181 CAGVGADYIKFLGISLKPFSYQVKGNYPLNVTMLFGGGYHKVGDHRRVERVEIAYHP 240
DB 181 CAGVGADYIKFLGISLKPFSYQVKGNYPLNVTMLFGGGYHKVGDHRRVERVEIAYHP 240
QY 241 TALSDVPRTTSASATLNTDYFGWIGRFRFAL 271
DB 241 TALSDVPRTTSASATLNTDYFGWIGRFRFAL 271

```



```
AC Q5HA08; Q5FCD7;
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Putative outer membrane protein Mapi-1 (Mapi-related protein).
GN Name=mapi-1; OrderedLocusNames=ERWE_CDS_09360, Erum8750;
OS Ehrlichia ruminantium (strain Welgevonden).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Ehrlichia.
OX NCBI_TaxID=254945;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15637156; DOI=10.1073/pnas.0406633102;
RA Collins N.E., Liebenberg J., de Villiers E.P., Brayton K.A., Louw B.,
RA Pretorius A., Faber F.E., van Heerden H., Josenmans A., van Kleef M.,
RA Steyn H.C., van Strijp M.F., Zweygardh E., Jongejan F., Maillard J.C.,
RA Berthier D., Botha M., Joubert F., Corton C.H., Thomson N.R.,
RA Allsopp M.T., Allsopp B.A.;
RT "The genome of the heartwater agent Ehrlichia ruminantium contains
RT multiple tandem repeats of actively variable copy number.";
RL Proc. Natl. Acad. Sci. U.S.A. 102:838-843(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Prutos R., Ferraz C., Bensaid A., Eychenie S., Kandassami Y., Viari A.,
RA Chantal I., Morgat A., Colisac E., Vachery N., Demaille J., Viari A.,
RA Martinez D.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR767821; CAH58610.1; -; Genomic DNA.
DR EMBL; CR925678; CA127420.1; -; Genomic DNA.
DR InterPro; IPR002566; Surface_Ag_msp4.
DR Pfam; PF01617; Surface_Ag_2; 1.
SQ Complete proteome.
SQ SEQUENCE 285 AA; 32108 MW; 05E4C4E59A6F9118 CRC64;

Query Match 62.2%; Score 892; DB 2; Length 285;
Best Local Similarity 59.1%; Pred. No. 4.5e-65;
Matches 166; Conservative 45; Mismatches 60; Indels 10; Gaps 4;

Qy 1 MNYKFFVGVVLTSTVF-FLSDGAFSDANFSEGRGLYIGSQYKVGIPNFSNFAEETIP 59
Db 5 MNYKEFVLGVTLSALLFSLPERAISDMVDSENRSRFYAGVQYRTGIPNFDNFSSETIP 64

Qy 60 GITKKIFALGLD--KSEINTHSNFTSRSDPTVYASSFAGSGIIGYVNDPFRPEGSYEN 117
Db 65 GLTKGVYGLDLKSDITKRAKNTFLYNTYSTSTGIGMGFYGFNIRMEPETSYS 124

Qy 118 FEPERQWYPENSQSKFFALSRN---ATNSP----NKFIVLENNGVVDKSLNVCYDIA 170
Db 125 FGIERQWYPEGQSQSHKFCVSRQNAAPTSSNNNDVFLNNGVKIRTLHVNFCYDVA 184

Qy 171 SGSIPLAPYMCAGVADYIKFLGISLPKFSYQVKFGVNYPLNVTMLFGGYYHKVVGDR 230
Db 185 HGNIPLAYVYVAGIGDVVKIGVSLPKFSYQLKFGVNYPLSIRTLMLFGGYYHKVWGSK 244

Qy 231 HERVEIAIHPALSDVPRTTSSASATLNTDYFGWEIGFRPAL 271
Db 245 YDRVKVYHPVQLNTVPKMTFVSANLIDIDYFGCEVGIRFFL 285

RESULT 8
Q4W4W3_EHRRU
ID Q4W4W3_EHRRU PRELIMINARY; PRT; 271 AA.
AC Q4W4W3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Mapi-1 (Fragment).
GN Name=mapi-1;
OS Ehrlichia ruminantium (Cowdria ruminantium).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Ehrlichia.
OX NCBI_TaxID=779;
RN [1]
```

```
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Senegal;
RX MEDLINE=22035368; PubMed=12039046; DOI=10.1016/S0378-1119(02)00408-0;
RA Bekker C.P., Bell-Sakyi L., Paxton E.A., Martinez D., Bensaid A.,
RA Jongejan F.;
RT "Transcriptional analysis of the major antigenic protein 1 multigene
RT family of Cowdria ruminantium.";
RL Gene 285:193-201(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Senegal;
RX PubMed=1595193; DOI=10.1128/JB.187.14.4782-4791.2005;
RA Bekker C.P., Postigo M., Taoufik A., Bell-Sakyi L., Ferraz C.,
RA Martinez D., Jongejan F.;
RT "Transcription Analysis of the Major Antigenic Protein 1 Multigene
RT Family of Three In Vitro-Cultured Ehrlichia ruminantium Isolates.";
RL J. Bacteriol. 187:4782-4791(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Senegal;
RA Bekker C.P.J., Taoufik A., Jongejan F.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF319940; AAY59059.1; -; Genomic DNA.
FT NON_TER 1
SQ SEQUENCE 271 AA; 30060 MW; F6AB82905DAA7680 CRC64;

Query Match 59.7%; Score 856; DB 2; Length 271;
Best Local Similarity 59.0%; Pred. No. 3.9e-62;
Matches 160; Conservative 42; Mismatches 61; Indels 8; Gaps 3;

Qy 9 GVTLSSTVF-FLSDGAFSDANFSEGRGLYIGSQYKVGIPNFSNFAEETIPGITKKIFA 67
Db 1 GVTLSALLFSLPERAISDMVDSENRSRFYAGVQYRTGIPNFDNFSASSETIPGLTKGVY 60

Qy 68 LGLD--KSEINTHSNFTSRSDPTVYASSFAGSGIIGYVNDPFRPEGSYENPEPERQWY 125
Db 61 LDLDLKSNGITKRAKNTFLYNTYSTSTGIGMGFYGFNIRMEPETSYSNFGVERQWY 120

Qy 126 PENSQSKFFALSRNAT----NSDNKFIVLNNGVVDKSLNVCYDIASGSIPLAPYM 180
Db 121 PEGSQSHKFCVSRNATPTSDSSNNNDVFLNNGVKVIRTLNVCYDVAHGIPLAPYV 180

Qy 181 CAGVGADYIKFLGISLPKFSYQVKFGVNYPLNVTMLFGGYYHKVVGDRHERVEIAIHP 240
Db 181 CAGIGGDYVYKFIGVSLPKFSYQLKFGVNYPLSIRTLMLFGGYYHKVYKSVYDRVKSLSS 240

Qy 241 TALSDVPRTTSSASATLNTDYFGWEIGFRPAL 271
Db 241 VQLSTVPKMTFVSANLIDIDYFGCEVGIRFFL 271

RESULT 9
Q5W5Y1_EHRRU
ID Q5W5Y1_EHRRU PRELIMINARY; PRT; 282 AA.
AC Q5W5Y1;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Mapi-1.
GN Name=mapi-1;
OS Ehrlichia ruminantium (Cowdria ruminantium).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Ehrlichia.
OX NCBI_TaxID=779;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Senegal;
RX MEDLINE=22035368; PubMed=12039046; DOI=10.1016/S0378-1119(02)00408-0;
RA Bekker C.P., Bell-Sakyi L., Paxton E.A., Martinez D., Bensaid A.,
RA Jongejan F.;
RT "Transcriptional analysis of the major antigenic protein 1 multigene
RT family of Cowdria ruminantium.";
RL Gene 285:193-201(2002).
```

```

RN [2] NUCLEOTIDE SEQUENCE.
RP STRAIN=Senegal;
RX PubMed=15995193; DOI=10.1128/JB.187.14.4782-4791.2005;
RA Bekker C.P., Postigo M., Taoufik A., Bell-Sakyi L., Ferraz C.,
RA Martinez D., Jongejan F.;
RT "Transcription Analysis of the Major Antigenic Protein 1 Multigene
RT Family of Three in Vitro-Cultured Ehrlichia ruminantium Isolates.";
RL J. Bacteriol. 187:4782-4791 (2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Senegal;
RA Bekker C.P.J., Taoufik A., Jongejan F.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF319940; AAV54088.2; -; Genomic DNA.
SQ SEQUENCE 282 AA; 31097 MW; 3D619510E7393EBB CRC64;

Query Match 36.6%; Score 525.5; DB 2; Length 282;
Best Local Similarity 41.8%; Pred. No. 6.1e-35;
Matches 120; Conservative 47; Mismatches 95; Indels 25; Gaps 10;

QY 1 MNYKFFVGVVTLSTVFVFLSDGAFSDANFS-----GRRGLYIGSQYKVGIPNFSNFA 54
DB 1 MNYKKILVRSALISLMSFLPYQSPAEPVSSNNIGNEXAKEGYISAKYNPSIPHRKESA 60

QY 55 EET----IPGITKKIFALGLDKSEINTHSNFTSRSDPTYASS---FAGSGIIGYYVNDF 107
DB 61 EETPVYKDSPTKKVFGCLKGGS-ITKYSDFTRT-DISFEGQNNFISGSGSIGYIMDGP 118

QY 108 RVEFGSYENPEPERQWYPEN----SQSYKFPALSRNATNSDNKFTVLENNGVVDKSLNV 163
DB 119 RVEIEAAVQKFNPKN---PANETDSDYYKHGSLRAETMTDKKVVLTNNGVTSSLMF 175

QY 164 NVCYDIASGSIPLAPYMCAGVADYIK-FLGISLPKFSYQVKGFGVNYPLNVNTMLFGGGY 222
DB 176 NACYDITAEVGFPIPYACAGIADLISIFDDINL-KFAYQKIGISYDITPEISAFIGY 234

QY 223 YHKVGVDRHERVEIAYHPTALSDVPRTTSASATLNTDYFGWEIGFRF 269
DB 235 YHGVIQNKYKVPVKL-PVTLTDAQSTASVTLTDAGYFGGELGVRF 280

RESULT 10
QSHA10 EHRRW PRELIMINARY; PRT; 282 AA.
AC QSHA10;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DE Putative outer membrane protein MAP1-1.
GN Name=map1-1; OrderedLocusNames=Erum8730;
OS Ehrlichia ruminantium (strain Welgevonden).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Ehrlichia.
OX NCBI_TaxID=254945;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15637156; DOI=10.1073/pnas.0406631102;
RA Collins N.E., Liebenberg J., de Villiers E.P., Brayton K.A., Louw E.,
RA Pretorius A., Faber P.E., van Heerden H., Josemans A., van Kleef M.,
RA Steyn H.C., van Strijp M.P., Zweggarth E., Jongejan F., Maillard J.C.,
RA Berthier D., Botha M., Joubert F., Corton C.H., Thomson N.R.,
RA Allsopp M.T., Allsopp B.A.;
RT "The genome of the heartwater agent Ehrlichia ruminantium contains
RT multiple tandem repeats of actively variable copy number.";
RL Proc. Natl. Acad. Sci. U.S.A. 102:838-843 (2005).
DR EMBL; CR767821; CAH58608.1; -; Genomic DNA.
DR InterPro; IPR011539; RHD.
DR InterPro; IPR002566; Surface Ag msp4.
DR InterPro; IPR011991; Wing hlx DNA_bd.
DR Pfam; PF01617; Surface_Ag_2; 1.
KW Complete proteome.
SQ SEQUENCE 282 AA; 31100 MW; 1CB2DA256B2CA990 CRC64;

```

```

Query Match 36.4%; Score 522.5; DB 2; Length 282;
Best Local Similarity 41.8%; Pred. No. 1.1e-34;
Matches 120; Conservative 47; Mismatches 95; Indels 25; Gaps 10;

QY 1 MNYKFFVGVVTLSTVFVFLSDGAFSDANFS-----EGRRGLYIGSQYKVGIPNFSNFA 54
DB 1 MNYKKILVRSALISLMSFLPYQSPAEPVSSNNIGNENAKEGYISAKYNPSIPHRKESA 60

QY 55 EET----IPGITKKIFALGLDKSEINTHSNFTSRSDPTYASS---FAGSGIIGYYVNDF 107
DB 61 EETPVYKDSPTKKVFGCLKGGS-ITKYSDFTRT-DISFEGQNNFISGSGSIGYIMDGP 118

QY 108 RVEFGSYENPEPERQWYPEN----SQSYKFPALSRNATNSDNKFTVLENNGVVDKSLNV 163
DB 119 RVEIEAAVQKFNPKN---PANETDSDYYKHGSLRAETMTDKKVVLTNNGVTSSLMF 175

QY 164 NVCYDIASGSIPLAPYMCAGVADYIK-FLGISLPKFSYQVKGFGVNYPLNVNTMLFGGGY 222
DB 176 NACYDITAEVGFPIPYACAGIADLISIFDDINL-KFAYQKIGISYDITPEISAFIGY 234

QY 223 YHKVGVDRHERVEIAYHPTALSDVPRTTSASATLNTDYFGWEIGFRF 269
DB 235 YHGVIQNKYKVPVKL-PVTLTDAQSTASVTLTDAGYFGGELGVRF 280

RESULT 11
QSHA10 EHRRW PRELIMINARY; PRT; 282 AA.
AC QSHA10;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Major antigenic protein 1 like protein (MAP1-1).
OS Ehrlichia ruminantium (Cowdria ruminantium).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Ehrlichia.
OX NCBI_TaxID=779;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99216274; PubMed=10198207; DOI=10.1006/bbrc.1999.0459;
RA Sulsona C.R., Mahan S.M., Barbet A.P.;
RT "The map1 gene of Cowdria ruminantium is a member of a multigene
RT family containing both conserved and variable genes.";
RL Biochem. Biophys. Res. Commun. 257:300-305 (1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Welgevonden;
RA van Heerden H., Collins N.E., Brayton K.A., Rademeyer C.,
RA Allsopp B.A.;
RT "Characterization of a major outer membrane protein multigene family
RT in Ehrlichia ruminantium.";
RL Gene 330:159-168 (2004).
DR EMBL; AF125276; AAD26347.1; -; Genomic DNA.
DR EMBL; AF125275; AAD26345.1; -; Genomic DNA.
DR EMBL; AY343331; AARI0944.1; -; Genomic DNA.
DR InterPro; IPR002566; Surface Ag msp4.
DR Pfam; PF01617; Surface_Ag_2; 1.
SQ SEQUENCE 282 AA; 31100 MW; 1CB2DA256B2CA990 CRC64;

Query Match 36.4%; Score 522.5; DB 2; Length 282;
Best Local Similarity 41.8%; Pred. No. 1.1e-34;
Matches 120; Conservative 47; Mismatches 95; Indels 25; Gaps 10;

QY 1 MNYKFFVGVVTLSTVFVFLSDGAFSDANFS-----EGRRGLYIGSQYKVGIPNFSNFA 54
DB 1 MNYKKILVRSALISLMSFLPYQSPAEPVSSNNIGNENAKEGYISAKYNPSIPHRKESA 60

QY 55 EET----IPGITKKIFALGLDKSEINTHSNFTSRSDPTYASS---FAGSGIIGYYVNDF 107
DB 61 EETPVYKDSPTKKVFGCLKGGS-ITKYSDFTRT-DISFEGQNNFISGSGSIGYIMDGP 118

QY 108 RVEFGSYENPEPERQWYPEN----SQSYKFPALSRNATNSDNKFTVLENNGVVDKSLNV 163

```





```
Db 61 EETPVYKDSPTKVKFGLKKEGS-ITKYSDFTRT-DISFEGQNNFISGFSIGYIMDGP 118
QY 108 RVEFEFGSYNFERQWYPEN-----SQSYKFPALSRNATNSDNKFIVLENNGVVDKSLNV 163
Db 119 RVEIEAAYQKFNPKN---PANETDTSYKYKHGLSRAEAMADKKYVVLTTNNGVTFSLSMF 175
QY 164 NVCYDIASGISPLAPYMCAGVADYIK-FLGISLPKFSYQVKFGVNYPLNVNTMLFGGGY 222
Db 176 NACYDITABGVPTIPYACAGIGADLISIFDDINL-KFAYQKGIGISYPTITPEISAFIGGY 234
QY 223 YHKVVGDRHERVEIAHYHPTALSVPRTTSASATLNTDYFGWEIGFRF 269
Db 235 YHGVIGNKYNKVPVKL-PVTLIDAPQSTSASVTLIDAGYFGGELGVRF 280
```

RESULT 15

```
QSFGL4_EHRRG
ID QSFGL4_EHRRG PRELIMINARY; PRT; 304 AA.
AC QSFGL4;
DT 10-MAY-2005 (Tremblrel. 30, Created)
DT 10-MAY-2005 (Tremblrel. 30, Last sequence update)
DT 10-MAY-2005 (Tremblrel. 30, Last annotation update)
DE Map1-related protein.
GN OrderedLocusNames=ERGA_CDS_09150;
OS Ehrlichia ruminantium (strain Gardel).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Anaplasmataceae; Ehrlichia.
OX NCBI_TaxID=302409;
RN [1]
NUCLEOTIDE SEQUENCE.
RC STRAIN=Gardel;
RA Frutos R., Ferraz C., Bensaid A., Eychenie S., Kandassami Y.,
RA Chantal I., Morgat A., Coissac E., Vachieri N., Demaille J., Viari A.,
RA Martinez D.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR925677; CAI28367.1; -; Genomic DNA.
DR InterPro; IPR002566; Surface_Ag_map4.
DR Pfam; PF01617; Surface_Ag_2; 1.
KW Complete proteome.
SQ SEQUENCE 304 AA; 33925 MW; 009338D2B65AE800 CRC64;
```

Query Match 35.9%; Score 514.5; DB 2; Length 304;  
Best Local Similarity 41.5%; Pred. No. 5.3e-34;  
Matches 119; Conservative 46; Mismatches 97; Indels 25; Gaps 10;

```
QY 1 MNYKRFVVGVTLSFVFLLSDGAFSDANFS-----EGRRGLYIGSQYKVGIPNFSNESA 54
Db 23 MNYKILVRSALISLSFYPQSFAFPVSSNNIGNENAKEGFYISAKYNPSIPHFKKFA 82
QY 55 EET-IPGITKKIPALGLDKSEINTHSNFRSDPTYASS---PAGFSGIIGYVNDP 107
Db 83 EETPVYKDSPTKVKFGLKKEGS-ITKYSDFTRT-DISFEGQNNFISGFSIGYIMDGP 140
QY 108 RVEFEFGSYNFERQWYPEN-----SQSYKFPALSRNATNSDNKFIVLENNGVVDKSLNV 163
Db 141 RVEIEAAYQKFNPKN---PANETDTSYKYKHGLSRAEAMADKKYVVLTTNNGVTFSLSMF 197
QY 164 NVCYDIASGISPLAPYMCAGVADYIK-FLGISLPKFSYQVKFGVNYPLNVNTMLFGGGY 222
Db 198 NACYDITABGVPTIPYACAGIGADLISIFDDINL-KFAYQKGIGISYPTITPEISAFIGGY 256
QY 223 YHKVVGDRHERVEIAHYHPTALSVPRTTSASATLNTDYFGWEIGFRF 269
Db 257 YHGVIGNKYNKVPVKL-PVTLIDAPQSTSASVTLIDAGYFGGELGVRF 302
```

Search completed: December 28, 2005, 15:38:00  
Job time : 114 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 28, 2005, 15:05:48 ; Search time 25 Seconds  
(without alignments)  
1042.990 Million cell updates/sec

Title: US-10-731-554-46  
Perfect score: 1434  
Sequence: 1 MNYKRFVGVVLTSTVFVFLS.....ASATLNTDYFGWEIGFRPAL 271

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 80.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	419.5	29.3	280	2 JE0217	28k surface antige
2	397.5	27.7	278	2 JE0216	28k surface antige
3	383.5	26.7	284	2 I40882	major antigenic pr
4	379	26.4	287	2 JE0220	28k surface antige
5	378.5	26.4	286	2 JE0219	28k surface antige
6	377	26.3	276	2 JE0218	28k surface antige
7	278	19.4	282	2 I39648	major surface prot
8	146.5	10.2	133	2 JE0221	28k surface antige
9	102	7.1	787	2 G64620	iron(III) dicitrat
10	100	7.0	284	2 AG3556	heat resistant agg
11	99	6.9	792	2 F71894	iron (iii) dicitra
12	97.5	6.8	614	2 E70120	hypothetical prote
13	97.5	6.8	713	2 F82506	probable TonB syst
14	97.5	6.8	991	2 B71315	conserved hypothet
15	96.5	6.7	1943	2 B64596	toxin-like outer m
16	96	6.7	730	2 A48832	cell surface glyco
17	95	6.6	1287	2 B53739	vacuolating cyto
18	93.5	6.5	3194	2 D71917	toxin-like outer m
19	93	6.5	624	2 E64483	hypothetical prote
20	93	6.5	1291	2 S44983	vacuolating cyto
21	92.5	6.5	637	2 F90257	hypothetical prote
22	92.5	6.5	726	2 T31887	hypothetical prote
23	92.5	6.5	2399	2 H71879	toxin-like outer m
24	91.5	6.4	239	2 AH0541	probable outer mem
25	91.5	6.4	460	2 T19515	hypothetical prote
26	91.5	6.4	1118	2 S75309	hypothetical prote
27	91	6.3	736	2 T12716	NADH2 dehydrogenas
28	90.5	6.3	336	2 B32808	flagellin chain B
29	90	6.3	237	2 S36343	opacity protein op

30 90 6.3 710 2 C98235 probable hydroxama  
31 90 6.3 716 2 A13050 hydroxamate-type f  
32 90 6.3 1101 2 S58108 hypothetical prote  
33 90 6.3 1769 2 S53378 probable membrane  
34 89 6.2 298 2 B71685 hypothetical prote  
35 89 6.2 482 1 A34671 triacylglycerol li  
36 89 6.2 493 2 G90604 hypothetical prote  
37 88.5 6.2 270 2 A71907 outer membrane pro  
38 88 6.1 797 2 D71621 hypothetical prote  
39 87.5 6.1 896 2 A51514 conserved membrane  
40 87.5 6.1 956 2 T40953 hypothetical prote  
41 87.5 6.1 1532 2 A26039 IGA-specific metal  
42 87.5 6.1 1561 2 S61314 IGA-specific metal  
43 87 6.1 374 2 I40765 hypothetical prote  
44 87 6.1 783 2 I50116 N-cadherin precurs  
45 87 6.1 937 2 S78561 CS3 piliin synthesi

ALIGNMENTS

RESULT 1

JE0217  
28k surface antigen 4 - Ehrlichia chaffensis  
N:Alternate names: MAP1  
C:Species: Ehrlichia chaffensis  
C>Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 09-Jul-2004  
C:Accession: JE0217  
R:Reddy, G.R.; Sulsana, C.R.; Barbet, A.F.; Mahan, S.M.; Burridge, M.J.; Alleman, A.R.  
Biochem. Biophys. Res. Commun. 247, 636-643, 1998  
A:Title: Molecular characterization of a 28kDa surface antigen gene family of the tribe  
A:Reference number: JE0216; MUID:98321180; PMID:9647746  
A:Accession: JE0217  
A:Molecule type: DNA  
A:Residues: 1-280 <RED>  
A:Cross-references: UNIPROT:O52107; UNIPARC:UPI000003478E; GB:AF062761

Query Match 29.3%; Score 419.5; DB 2; Length 280;  
Best Local Similarity 34.1%; Pred. No. 1.2e-27;  
Matches 98; Conservative 59; Mismatches 103; Indels 27; Gaps 7;

QY 1 MNYKRFVGVVLTSTVFVFLSDGAFSDANFSEGRRG-LVIGSOYKVGIPNFSNFAETIP 59  
DB 1 MNCKRFFITTLVLSMSFLPGISFSDAVQNDNVGNFVIGSKYVPSVSHFGVSAKQE-R 59  
QY 60 GITKKIFAL--GLDKSEINTHSNFRSYDPTA-----SSPAGFSGIIGYYVNDPRVEPE 112  
DB 60 NTTIGVFLGKQWDGSGTISKNSPENTVNPVNSFKYNNPFLGFAGAVCYLMNGPRIELE 119  
QY 113 GSYENFEPERQWYPSNSQSYKFFALSRNA----TNSDNKFIVLENNGVVDKSLNVNVCYD 168  
DB 120 MSYETFDVKNQGNKYKNDAAHYALTHNSGGKLSNAGDKFVFLKNEGLLDISIMLNACYD 179  
QY 169 IASGSIPLAPYMCAGVADYIKFLGSLPKFSYQVKGWYPLNVTNMLFGGYVHKVVG 228  
DB 180 VISEGIPSPYICAGVGTDLISMPFAINPKISYQKGLSISISPEASVFGVGHFKVIG 239  
QY 229 DRHERVEIAYHPTALSDVPRITSAS-----ATLNTDYFGWEIGFRF 269  
DB 240 NEFRDIP-----AMIPSTSLTGNHFTIVTLVSVCHFGVELGGRF 278

RESULT 2

JE0216  
28k surface antigen 3 - Ehrlichia chaffensis  
N:Alternate names: MAP1  
C:Species: Ehrlichia chaffensis  
C>Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 09-Jul-2004  
C:Accession: JE0216  
R:Reddy, G.R.; Sulsana, C.R.; Barbet, A.F.; Mahan, S.M.; Burridge, M.J.; Alleman, A.R.  
Biochem. Biophys. Res. Commun. 247, 636-643, 1998  
A:Title: Molecular characterization of a 28kDa surface antigen gene family of the tribe  
A:Reference number: JE0216; MUID:98321180; PMID:9647746

A:Accession: JE0216  
A:Molecule type: DNA  
A:Residues: 1-278 <RED>  
A:Cross-references: UNIPROT:052106; UNIPARC:UPI000003478D; GB:AF062761

Query Match 27.7%; Score 397.5; DB 2; Length 278;  
Best Local Similarity 34.7%; Pred. No. 8.1e-26;  
Matches 99; Conservative 55; Mismatches 106; Indels 25; Gaps 9;

Qy 1 MNYKRFVGVVTLSTFVFLSDGAFSDANFSEGRRLG-LYIGSQYKVGIPNFSNFSABETI 58  
Db 1 MNCKKFFITTTALVLSLFLGSLFSDVPQGNISGNFVSGKMPASASHFGMFSKAEKN 60

Qy 59 PGITKKIFAL-----GLDKSEINTHSNFTSYDPTYASS-FAGFSGLIIGYVNDFRVEFE 112  
Db 61 P--TVALYGLQDMEGISSSHNDHFNKGVSKFYENNPFGLPAGALGYSGMGGRPRVEFE 118

Qy 113 GSYENFEPERQWYPENSQYKFFALSRNATN---SDNKFIVLENNGVVDKSLNVNVCYDI 169  
Db 119 VSYETFDVKNQGNKYKNDHRYCALGQODNSGIPKTSKYVLKSEGLLDISFMLNACVDI 178

Qy 170 ASGSTPLAPYMCAGVADYIKFLGSLPKFSYQVKFGVNYPLNVNMTLFGGYYHKVGVGD 229  
Db 179 INESLPLSPYICAGVGTDLISMFENATPKISTQGKLGUSYINPEASVFIGGHFKVIGN 238

Qy 230 RHERVEIAYHPTALSDVPRITTSAS-----ATLNTDYFGWEIGGRF 269  
Db 239 EFRDI-----PTLKAFV--TSSATPDLAIVLSVCHFGIELGGRF 276

RESULT 3  
140882  
major antigenic protein - heartwater rickettsia  
C:Species: Cowdria ruminantium (heartwater rickettsia)  
C:Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 09-Jul-2004  
C:Accession: I40882; S42827  
R:van Vliet, A.H.; Jongejan, F.; van Kleef, M.; van der Zeijst, B.A.  
Infect. Immun. 62, 1451-1456, 1994  
A:Title: Molecular cloning, sequence analysis, and expression of the i  
A:Reference number: I40882; MUID:94178956; PMID:8132352  
A:Accession: I40882  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-284 <RES>  
A:Cross-references: UNIPROT:Q46327; UNIPARC:UPI000009A44; EMBL:X74250; NID:G454266; PID  
C:Genetics:  
A:Gene: map1

Query Match 26.7%; Score 383.5; DB 2; Length 284;  
Best Local Similarity 32.9%; Pred. No. 1.2e-24;  
Matches 93; Conservative 55; Mismatches 120; Indels 15; Gaps 7;

Qy 1 MNYKRFVGVVTLSTFVFLSDGAFSDANFSEGR--RGLYIGSQYKVGIPNFSNFSABETI 58  
Db 1 MNCKKIFITTLISLVSFLPGVSFSDVTEENNPVGSVYISAKYMTPTASHFGMFSKAEKS 60

Qy 59 PGITKKIFAL-----GLDKSEINTHSNFT-RSYDPTYASS-FAGFSGLIIGYVNDFRVEFE 111  
Db 61 RD-TKAVFLGKKDWDGVKTPSGNTNSITFEKDYSEKYNPNFPLGAGAVGSMGPRIEP 119

Qy 112 EGSYENFEPERQW--YPENSQYKFFALSRNATNSDNKFIIVLENNGVVDKSLNVNVCYDI 169  
Db 120 EVSYETFDVRNPGGNYKNDHRYCALDHTASSSTAGATTSVMVKENLTDISLMLNACVDI 179

Qy 170 ASGSTPLAPYMCAGVADYIKFLGSLPKFSYQVKFGVNYPLNVNMTLFGGYYHKVGVGD 229  
Db 180 MLDGMPVSPYVCAGIGTDLVSVINATNPKLSYQKLGISYSINPEASVFIGGHFHRVIGN 239

Qy 230 RHERVEIAYHPTALSDVPRITTS-----ASATLNTDYFGWEIGGRF 269  
Db 240 EFKDIATSKVTSSGNASSAVSPGFASAILDVCHFGIELGGRF 282

RESULT 4  
JE0220  
28k surface antigen 1 - Ehrlichia canis  
C:Species: Ehrlichia canis  
C:Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 09-Jul-2004  
C:Accession: JE0220  
R:Reddy, G.R.; Sulsona, C.R.; Barbet, A.F.; Mahan, S.M.; Burridge, M.J.; Alleman, A.R.  
Biochem. Biophys. Res. Commun. 247, 636-643, 1998  
A:Title: Molecular characterization of a 28kDa surface antigen gene family of the tribe  
A:Reference number: JE0216; MUID:98321180; PMID:9647746  
A:Accession: JE0220  
A:Molecule type: DNA  
A:Residues: 1-287 <RED>  
A:Cross-references: UNIPROT:Q9ZGJ0; UNIPARC:UPI0000034790; GB:AF062762; NID:G3327964; PID

Query Match 26.4%; Score 379; DB 2; Length 287;  
Best Local Similarity 34.1%; Pred. No. 3e-24;  
Matches 101; Conservative 55; Mismatches 102; Indels 38; Gaps 12;

Qy 1 MNYKR-FVVG--VTLSTFVFLSDGAFSDANFSEGRRLG-LYIGSQYKVGIPNFSNFSABET 57  
Db 1 MNYKFTVTVALVLLTSFTFIP--FVSPARST-IHNFYISGKYMPTASHFGIFSAAKEE 57

Qy 58 IPGITKKIFALGDK-----SEINTH-----SNFRSYDPTYASS-FAGFSGLIIGYVNDFR 108  
Db 58 -QSFTKVL--VGLDQRLSHNIIINNNDTAKSLKVQNSFKYKNNPFLGFAIGAIGYSIGNSR 114

Qy 109 VFEFGSYENFEPERQWYPENSQYKFFALSRNA-----TNSDNKFIIVLENNGV 156  
Db 115 IELEVSHEIFDTRKNPGNLYNDSHKCALSHGSHICSDGNSGDWYTAKTDFVLLKNEGL 174

Qy 157 VDKSLNVNVCYDIASGSIPLAPYMCAGVADYIKFLGSLPKFSYQVKFGVNYPLNVNMT 216  
Db 175 LDVSPMLNACYDITTEKMPSPYICAGIGTDLISMETTQNKISYQKGLGNTINSRVS 234

Qy 217 LFGGYYHKVGVDRHERVEIAYHPTAL---SDVPRITTSASATLNTDYFGWEIGGRF 269  
Db 235 VFAGGHFKVIGNEFKGI-----PTLLPDGNSIKVQOSATVTLDVCHFGLEIGSRF 285

RESULT 5  
JE0219  
28k surface antigen 2 - Ehrlichia chaffeensis  
N:Alternate names: MAP1  
C:Species: Ehrlichia chaffeensis  
C:Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 09-Jul-2004  
C:Accession: JE0219  
R:Reddy, G.R.; Sulsona, C.R.; Barbet, A.F.; Mahan, S.M.; Burridge, M.J.; Alleman, A.R.  
Biochem. Biophys. Res. Commun. 247, 636-643, 1998  
A:Title: Molecular characterization of a 28kDa surface antigen gene family of the tribe  
A:Reference number: JE0216; MUID:98321180; PMID:9647746  
A:Accession: JE0219  
A:Molecule type: DNA  
A:Residues: 1-286 <RED>  
A:Cross-references: UNIPROT:O52105; UNIPARC:UPI000003478C; GB:AF062761

Query Match 26.4%; Score 378.5; DB 2; Length 286;  
Best Local Similarity 31.4%; Pred. No. 3.3e-24;  
Matches 93; Conservative 58; Mismatches 104; Indels 41; Gaps 9;

Qy 1 MNYKRFVGVVTLSTFVFLSDGAFSDANFSEGRRLG-LYIGSQYKVGIPNFSNFSABETIP 59  
Db 1 MNCKKFFITTTALTLMSFLPGISLSDVPQDNDISGNFYISGKYMPSASHFGVFSKAEK-R 59

Qy 60 GITKKIFALGDL-----KSEINTHSNFTSYDPTYASSFAGFSGLIIGYVNDFRV 109  
Db 60 NTTVGVFGBQDWRCDRCVISRTTSLSDIFTVPNYSFKYE---NNLFSGFAIGAIGYSMDGPR 116

Qy 110 EFGSYENFEPERQWYPENSQYKFFALSR-----NATNSDNKFIIVLENNGVVDKSL 161  
Db 117 ELEVSHEAFDVKQGNKYKNEAHRYALSHLLTETQIDGAGSASVPLI--NEGILLDKSP 174

Qy 162 NVNVCYDIASGSIPLAPYMCAGVADYIKFLGSLPKFSYQVKFGVNYPLNVNMTLFGGG 221

```
Db      175 MLNACYDVISGIFSPYICAGIGIDLVSFPEANPKISYQKGLGYSIPSPASVFTGG 234
Qy      222 YHKVVGDRHERVEIAYHPTALSDVPRRTS-----ASATLNTDYFGWEIGFRF 269
Db      235 HPHKVIQNEPRDI-----PTM---IPSSALAGKNYPAIVLTVFFYFIELGGRF 282

RESULT 6
JB0218
28k surface antigen 5 - Ehrlichia chaffensis
N:Alternate names: MAP1
C:Species: Ehrlichia chaffensis
C:Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 09-Jul-2004
C:Accession: JB0218
R:Reddy, G.R.; Sulsona, C.R.; Barbet, A.F.; Mahan, S.M.; Burridge, M.J.; Alleman, A.R.
Biochem. Biophys. Res. Commun. 247, 636-643, 1998
A:Title: Molecular characterization of a 28kDa surface antigen gene family of the tribe
A:Reference number: JB0216; MUID:98321180; PMID:9647746
A:Accession: JB0218
A:Molecule type: DNA
A:Residues: 1-276 <RED>
A:Cross-references: UNIPROT:Q93DD2; UNIPROT:Q9ACI9; UNIPARC:UPI000003478F; GB:AF062761

Query Match      26.3%; Score 377; DB 2; Length 276;
Best Local Similarity 29.8%; Pred. No. 4.2e-24;
Matches 87; Conservative 57; Mismatches 106; Indels 42; Gaps 6;

Qy      1 MNYKRFVGVGTLSTVFVFLSDGAFSDANFSEGRGLGYSQYKVGIPNFSNFSABETIPG 60
Db      1 MNYKVFITSLISLISLPGVFSFDGAGNGNFYISGKYMPSASHFGVFSABEE-RN 59

Qy      61 ITKKIFAL-----GLDKSEINTHSNFTSRSDYTYASSFAGSGIIGYVNDFRVE 110
Db      60 TTGVGFLKQNWGSAINSNSPNDVTVSNYSFKYE---NNPFLGAGAGYSMDGPRIE 116

Qy      111 FPGSYENFEPERQWTPENSQSYKFFALSRNA-----TNSDNKFTVLNNGVVDKSLNVVC 166
Db      117 LEVSYETFDVKQNGNKNYNEAHRYCALSHNSAADSSASNNFVFLNKGGLDISFMLNAC 176

Qy      167 YDIASGSIPLAPYMCAGVADYIKFLGISLPKFSYQVKGFGVNYPLNVTMLFGGGYHYKV 226
Db      177 YDVGEGIPFPYSICAGITGLVSMFEATNPKISYQKGLGYSIPSPASVFTGGHFHKV 236

Qy      227 VGDHRERVEIAYHPTALSDVPRRTSASAT-----INTDYFGWEIG 266
Db      237 IGNE-----FRDPIITPGSTLAGKNYPAIVLTVDFCHFGIENG 276

RESULT 7
I39648
major surface protein 4 - Anaplasma marginale
C:Species: Anaplasma marginale
C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 09-Jul-2004
C:Accession: I39648
R:Oberle, S.M.; Barbet, A.F.
Gene 136, 291-294, 1993
A:Title: Derivation of the complete map4 gene sequence of Anaplasma marginale without cl
A:Reference number: I39648; MUID:94124017; PMID:8294020
A:Accession: I39648
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-282 <RES>
A:Cross-references: UNIPROT:Q07408; UNIPARC:UPI0000000B7F; GB:L01987; NID:gi42174; PIDN:
C:Genetics:
A:Gene: map4

Query Match      19.4%; Score 278; DB 2; Length 282;
Best Local Similarity 26.1%; Pred. No. 8.2e-16;
Matches 77; Conservative 57; Mismatches 119; Indels 42; Gaps 10;

Qy      1 MNYKR-FVGVGTLSTF--VFFISDGA-----FSDANFSEG---RRGLYIGSQYKVGIPNF 49
```

```
Db      1 MNYRELFTGCLSAATVACCSLLVSGAVVASPMSEHVASEGVGMGGSFVYVCAAYSPAPFSV 60
Qy      50 SNFSAEETIPGITTKIFALGLDKS-----EINTHSNFTSRSDYTYASS---FAGSGIIGY 102
Db      61 TSDMRES---SKETSIVRGYDKSIATIDVSVPAFNSKS-GYTFAPSKNLTISFDGAVGY 116
Qy      103 YVNDPRVFEGSYENFEPERQWTPENSQSYKFFALSRNATNSDNKFTVLNNGVVDKSLN 162
Db      117 SLGGARVELEASRYRRFATLADGQYAKSGAESLAAITRANITETNYFVVKIDEITNTSVM 176
Qy      163 VNVYDIASGSIPLAPYMCAGVADYIKFLGISLPKFSYQVKGFGVNYPLNVTMLFGGGY 222
Db      177 LNCYDVLHTDLVSPVVCAGIGASFDVISKQVTTKLAYRGKVGISYQFTPEISLVAGGF 236
Qy      223 YHKVVGDRHERVEIAYHPTALSDVPRRTS-----ASATLNTDYFGWEIGFRF 269
Db      237 YHGLFDESRYK-----DIPAHNSVKFSGEAKASVKAHTIADYGFNLGARF 279

RESULT 8
JB0221
28k surface antigen 2 - Ehrlichia canis
C:Species: Ehrlichia canis
C:Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 09-Jul-2004
C:Accession: JB0221
R:Reddy, G.R.; Sulsona, C.R.; Barbet, A.F.; Mahan, S.M.; Burridge, M.J.; Alleman, A.R.
Biochem. Biophys. Res. Commun. 247, 636-643, 1998
A:Title: Molecular characterization of a 28kDa surface antigen gene family of the tribe
A:Reference number: JB0216; MUID:98321180; PMID:9647746
A:Accession: JB0221
A:Molecule type: DNA
A:Residues: 1-133 <RED>
A:Cross-references: UNIPROT:O85360; UNIPARC:UPI00000034791; GB:AF062762; NID:93327964; PI

Query Match      10.2%; Score 146.5; DB 2; Length 133;
Best Local Similarity 33.6%; Pred. No. 3e-05;
Matches 45; Conservative 23; Mismatches 51; Indels 15; Gaps 7;

Qy      1 MNYKR-FVGVGTLSTVFVFLSDGAFSDANFSEGRGLGYSQYKVGIPNFSNFSABETI 58
Db      1 MNCKKVFITSLISS-IVFLPNVSYSNPVVYSGNSMYGNFYISGKYMPSVPHFGIFSABEE- 58

Qy      59 PGITTKIFALGLDK-----SEINTHSNFT-RSDYTYASS-FAGSGIIGYVNDPRV 109
Db      59 --KKKTTVVYGLKENWAGDAISSQSPDNFTIRNYSFKYASNKFLGFAVAIGSIGSPRI 116

Qy      110 BFEGSYENFEPERQ 123
Db      117 EVMSYEAFVKNQ 130

RESULT 9
G64620
iron(III) dicitrate transport protein - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C:Accession: G64620
R:Tomb, J.F.; White, O.; Kervlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Matthey, L.
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: G64620
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-787 <TOM>
A:Cross-references: UNIPROT:O25487; UNIPARC:UPI000000D3029; GB:AE000592; GB:AE000511; NID
C:Superfamily: ferrichrome-iron receptor; tonB-dependent receptor amino-terminal homolog
C:Keywords: membrane protein

Query Match      7.1%; Score 102; DB 2; Length 787;
```



Db 249 WYFIGKLSLQDDF---FEKS-----QRMQLAKNGSIYLTAYTLRNKKAVERKFEIK 299  
QY 153 NNG-----VVD-KSLNVNVCY 167  
Db 300 DSGMNAVVIDFKDDGNLTYY 319

RESULT 13  
F82506  
probable TonB system receptor VCA0064 [imported] - Vibrio cholerae (strain N16961 serogroup O139)  
C:Species: Vibrio cholerae  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C:Accession: F82506  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Peterson, S.; Loftholm, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, J.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833; PMID:10952301  
A:Accession: F82506  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-713 <HEI>  
A:Cross-references: UNIPROT:Q9KNA1; UNIPARC:UPI00000333F5; GB:AE004349; GB:AE003853; NID  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VCA0064  
A:Map position: 2

Query Match 6.8%; Score 97.5; DB 2; Length 713;  
Best Local Similarity 21.2%; Pred. No. 3.5;  
Matches 65; Conservative 36; Mismatches 116; Indels 89; Gaps 14;

QY 21 DGAFSDANFSEGRGLYIGSVKVGIPNFSN-----PSAETIPGITK---- 63  
Db 388 DLQADEPPAPAR-----EYFGVGDNLLGDWTLAIGRDAQKLSKPNTRHG 440

QY 64 -KIFALGLDK-----SEINTNSNFTN-----SYDPTASSFAGFGIIGYVN 105  
Db 441 YKVVTMGSEWSPSASISYQHPWNTYLSYNGHFRAPSYKAYGASDHSPVLPFTII- 499

QY 106 DFRVEPEGSYENFEPERQWYKPFALSRNATNSDNKFIVLNNGVVDKSLNV-N 164  
Db 500 -----KPNKLAETSDS---FELGSKYDNGQTQFYAVAFYSIFDNFDVKQ 543

QY 165 VCYDIASGSIPLAPYMCAGVADYIKFLGISLP-----KFSY-QVKFGVNP 210  
Db 544 VGYDNATGSIQYQYQNIAGV-----KTYGAEMSVHRLDDRWSVENKLGVDGKGENY 598

QY 211 LNVNTMLFGGGY--YHKVVGDRHERVEIAYHPTALSDVPRITS-----ASATLNTDYFGW 263  
Db 599 VRTLPLEGVSQVQLYQERWDAYSRLNWA---SAMSRSVPTCTEKGKETECATTGGWSM 655

QY 264 BIGFRF 269  
Db 656 DIGLNY 661

RESULT 14  
B71315  
conserved hypothetical protein TP0515 - syphilis spirochete  
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: B71315  
R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwinn, M.L.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDermott, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
Science 281, 375-388, 1998  
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
A:Reference number: A71250; MUID:98332770; PMID:9665876  
A:Accession: B71315  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA  
A:Residues: 1-991 <COL>  
A:Cross-references: UNIPROT:O83528; UNIPARC:UPI000000COA89; GB:AE001227; GB:AE000520; NID  
A:Experimental source: strain Nichols  
C:Genetics:  
A:Gene: TP0515

Query Match 6.8%; Score 97.5; DB 2; Length 991;  
Best Local Similarity 22.0%; Pred. No. 5.4;  
Matches 51; Conservative 36; Mismatches 94; Indels 51; Gaps 9;

QY 42 YKVGIPNFSASSETIPGITKIFALGLDKSEINTNSNFTNTRSYDPTVASSFAGFGIIG 101  
Db 723 YAYSIAITFLSP-----PYISATLSARFLDPQVRAELDFRMPYEITCKQT----- 767

QY 102 YYVNDFRVEPEGSYENFEPERQWYKPFALSRNATNSDNKFIVLNNGVVDKSL 161  
Db 768 -YVYDIGKMSDSY-----EVSGLGWKPYLSYLLKGESGKAAPSGNNGLSISKL 815

QY 162 NVNVCYD-----IASGSIPLAPYMCAGVADYIKFLGISLPKFSYQVKFGVN--YPLN 212  
Db 816 KLTLSHDDPFLTVRFWKRRIKIQGTLSLSLEINF-KLDKSHISFSPDLITTSYKFLDLS 874

QY 213 VNTMLFGGGYVHKVVGDRHERVEIAYHPTALSDVPRITSASATLNTDYFGWE 264  
Db 875 INTVI-----KNEKL-TPYFPTQSSQSPOTKLWDAFVSSLYP-WD 912

RESULT 15  
B64596  
toxin-like outer membrane protein HP0610 - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 09-Jul-2004  
C:Accession: B64596  
R:Tomb, J.F.; White, O.; Kexlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenna, J.; Venter, J.C.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Matthey, L.; Son, J.D.; Kelley, J.M.; 1997  
Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467; PMID:9252185  
A:Accession: B64596  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1943 <TOM>  
A:Cross-references: UNIPROT:O25331; UNIPARC:UPI00000D30AD; GB:AE000575; GB:AE000511; NID  
C:Genetics:  
A:Start codon: GTG

Query Match 6.7%; Score 96.5; DB 2; Length 1943;  
Best Local Similarity 22.9%; Pred. No. 16;  
Matches 62; Conservative 42; Mismatches 102; Indels 65; Gaps 16;

QY 1 MNYKRFVVGVTI-----STFVFFLSDGAFSDANFSEGRGLYIGSVKVGIPN 48  
Db 85 LNFQATVVGVTLLLSAQKANNNGSIWFGKNNLLYLHGNFN--ATNIFLTNNFVGNPN 142

QY 49 FS-----NFSABETI--PGITKKIF---ALGLDKSEINTNS--NFTRSYDPTYA--SSFAG 95  
Db 143 AGCGATINFNADETLDNADGLNVTNFQTVALGLQTS-ASQHSWANFNSKLSMEIKNSNFRD 201

QY 96 FS-GIIGYVNDFRVEFE-----GSYENPEPERQWYKPFALSRNATNSDNKFIVLNNGVVDKSLNVNVCYDIASGSIPLAPYMCAGV--OSQYK 134  
Db 202 FTWG--GFNFSGRITFTENTFTSGWTNNGATGESGYSYNNVNVANTDIFNSILGGIRY 259

QY 135 FALSBNATNSDNKFIVLNNGVVDKSLNVNVCYDIASGSIPLAPYMCAGV--OSQYK 185  
Db 260 DLKANNIIFNNSQWVIDVSKVNSQSLNGNVTNF--NSRLSVKPNNAINIGDSQQTQTALE 317

QY 186 -ADYIKFLGISLPKFSYQVKF-GVNYPLNVN 214  
Db 318 NASSLSFYNNVANSVANGTATFNGVSY-LNLN 347

Search completed: December 28, 2005, 15:38:31  
Job time : 27 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 28, 2005, 15:38:40 ; Search time 18 Seconds  
(without alignments)  
112.748 Million cell updates/sec

Title: US-10-731-554-46  
Perfect score: 1434  
Sequence: 1 MNVRFVGVGVTSTFVFLS.....ASATLNTDYFGWEIGFRPAL 271

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 57103

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA\_New.\*  
1: /cgn2\_6/ptodata/1/pubaa/US08\_NEW\_PUB pep.\*  
2: /cgn2\_6/ptodata/1/pubaa/US06\_NEW\_PUB pep.\*  
3: /cgn2\_6/ptodata/1/pubaa/US07\_NEW\_PUB pep.\*  
4: /cgn2\_6/ptodata/1/pubaa/PCT\_NEW\_PUB pep.\*  
5: /cgn2\_6/ptodata/1/pubaa/US09\_NEW\_PUB pep.\*  
6: /cgn2\_6/ptodata/1/pubaa/US10\_NEW\_PUB pep.\*  
7: /cgn2\_6/ptodata/1/pubaa/US11\_NEW\_PUB pep.\*  
8: /cgn2\_6/ptodata/1/pubaa/US60\_NEW\_PUB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	5.9	1565	6	US-10-467-657-2704
2	84	5.9	477	7	US-11-055-822-98
3	83.5	5.8	1433	7	US-11-094-519A-40
4	81	5.6	876	7	US-11-077-550-128
5	81	5.6	877	7	US-11-077-550-126
6	81	5.6	877	7	US-11-077-550-130
7	81	5.6	881	7	US-11-077-550-124
8	81	5.6	902	7	US-11-077-550-132
9	81	5.6	912	7	US-11-077-550-116
10	81	5.6	914	7	US-11-077-550-120
11	81	5.6	944	7	US-11-077-550-122
12	81	5.6	950	7	US-11-077-550-118
13	78	5.4	219	7	US-11-170-653-29
14	77.5	5.4	1122	7	US-11-191-374-3
15	77.5	5.4	1122	7	US-11-191-375-3
16	77.5	5.4	1122	7	US-11-191-588-3
17	76	5.3	1296	6	US-10-615-668-3
18	76	5.3	1468	6	US-10-467-657-1088
19	76	5.3	2053	7	US-11-013-759-9
20	75	5.2	1992	7	US-11-013-759-3
21	75	5.2	1992	7	US-11-013-759-13
22	75	5.2	2047	7	US-11-013-759-4
23	75	5.2	2047	7	US-11-013-759-7
24	74	5.2	826	6	US-10-793-626-1066
25	73	5.1	278	6	US-10-467-657-7386

26	72.5	5.1	265	6	US-10-467-657-1778	Sequence 1778, Ap
27	72	5.0	475	6	US-10-510-386-98	Sequence 98, Appl
28	71.5	5.0	173	6	US-10-984-376-13	Sequence 13, Appl
29	71.5	5.0	408	6	US-10-793-626-2286	Sequence 2286, Ap
30	71.5	5.0	527	7	US-11-120-543-18	Sequence 18, Appl
31	71	5.0	497	7	US-11-037-829A-15	Sequence 15, Appl
32	70.5	4.9	174	6	US-10-984-376-10	Sequence 10, Appl
33	70	4.9	222	7	US-11-170-653-32	Sequence 32, Appl
34	70	4.9	223	7	US-11-170-653-30	Sequence 30, Appl
35	70	4.9	223	7	US-11-170-653-31	Sequence 31, Appl
36	69.5	4.8	174	6	US-10-984-376-7	Sequence 7, Appl
37	69.5	4.8	174	6	US-10-984-376-8	Sequence 8, Appl
38	69.5	4.8	221	7	US-11-170-653-44	Sequence 44, Appl
39	69.5	4.8	428	7	US-11-074-176-364	Sequence 364, App
40	69	4.8	175	6	US-10-467-657-2512	Sequence 2512, Ap
41	69	4.8	190	7	US-11-170-653-33	Sequence 33, Appl
42	69	4.8	381	6	US-10-793-626-184	Sequence 184, App
43	68.5	4.8	174	7	US-11-103-957-90	Sequence 90, Appl
44	68.5	4.8	1259	6	US-10-467-657-5510	Sequence 5510, Ap
45	68.5	4.8	1438	6	US-10-511-559-73	Sequence 73, Appl

ALIGNMENTS

RESULT 1  
US-10-467-657-2704  
; Sequence 2704, Application US/10467657  
; Publication No. US20050260581A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON Spa  
; APPLICANT: FONTANA Maria Rita  
; APPLICANT: PIZZA Mariagrazia  
; APPLICANT: MASIGNANI Vega  
; APPLICANT: MONACI Elisabetta  
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/467,657  
; CURRENT FILING DATE: 2003-08-11  
; PRIOR APPLICATION NUMBER: GB-0103424.8  
; PRIOR FILING DATE: 2001-02-12  
; NUMBER OF SEQ ID NOS: 9218  
; SOFTWARE: Seqwin99, version 1.04  
; SEQ ID NO 2704  
; LENGTH: 1565  
; TYPE: PRT  
; ORGANISM: Neisseria gonorrhoeae  
US-10-467-657-2704

Query Match	5.9%	Score 85;	DB 6;	Length 1565;
Best Local Similarity	21.7%	Pred. No. 3.9;		
Matches	60;	Conservative 37;	Mismatches 96;	Indels 84; Gaps 15;
Qy	2	NYKRFVGVGT-LSTFVFFLSDGAFSDANFSEGRGLVIGSOY-KVGIPNFSNFSAEETIP	59	
Db	17	NKGRFFVGATDLV-----KNKRGQGNALSNVPMDFSVADNKRKA	60	
Qy	60	GITKKIFALGLDKSEINTHSNFTSRSDPTVASSPAGFSGIIGYYVNDPFRVEFGSYENFE	119	
Db	61	TVVDPQYAVSVKKAKEVHT-----FYYQYNGHNDVAD-KENEYRVV---EQNNYE	108	
Qy	120	PERQWYPENSQSYKFFALSBNATNSDNKFIVLNNGVVDKSLNVNVCYDTASGSIPLAPY	179	
Db	109	PHKAWASNLGRLEDYNNAR-----FNKFVT-----EVA-----PIAP-	141	
Qy	180	MCAGVGADYIKFLGISLPKFSYQVKFGVNYPLNVNTMLFGGGYVHKVVDGDSHVEVAYH	239	
Db	142	TDAGGGGLDTYK-----DKNRFSFVRIGAG-----RQLVKEGVYHQ-----EGNEKGYD	186	
Qy	240	PTALSDVPRTTASA-----TLNTDYFGWEIGF	267	
Db	187	LRDLQAYRAIAGTPYKDNIDQTMTEGL---IGF	220	



; NUMBER OF SEQ ID NOS: 179  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 128  
; LENGTH: 876  
; TYPE: PRT  
; ORGANISM: Clostridium botulinum  
US-11-077-550-128

Query Match            5.6%; Score 81; DB 7; Length 876;  
Best Local Similarity 20.4%; Pred. No. 4.4;  
Matches 68; Conservative 39; Mismatches 100; Indels 126; Gaps 16;

QY         5 RFVVGVTLSFTVFFLSDGAFSDANF-----SEGRRLGLYGISQYKVGIPNFSNFAE 55  
            :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
DB         208 RPM--LTYSNATNDVGEGRFSKFECMDPILIMHMLNHAMH--NLGYIAIPN-----D 257

QY         56 ETIPGITKKIF----ALGLDKSEINTHSNTFRSYDPTYASSF-----93  
            :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
DB         258 QTISSVTSNIFFSQYNVKLEYAEIYAFGGPTIDILPKSARKYPFEKALDYRSIAKLNS 317

QY         94 -----AGPSGIIGYY-----VNDRFVEFGSYE---NFEPERQWYPENSQYKFFALS RN 140  
            :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
DB         318 ITTANPSSFNKYIGEYKOKLRKYRFVVESSEGVTVNRNKPFVELYNELTQTTFEFNYAK- 376

QY         141 ATNSDNKFIULEN--NGVVDKSLNVNCYDIASGISPLAPYMCAGVADYIKFLGIGSLPK 198  
            :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
DB         377 IYNQVRKIYLSNVVTPVTANILDNV-YDIQN-----GFNIPK 413

QY         199 FSQVQKF-GVNY-----PLANVTMLF-----GGGYHKVVG-----228  
            :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
DB         415 SNLVLFMGQLSRNPALKRVKNPENMLYLFTFKCHKAIDGRSLYNTKDLCRELLVKNTDL 474

QY         229 -----DRHERVEIAHYHTALS 244  
            :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
DB         475 PFIGDISDVKTDFILRKDINEETEVIYYPDNV 507

RESULT 6  
US-11-077-550-130  
; Sequence 130, Application US/11077550  
; Publication No. US2005024435A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford Charles  
; APPLICANT: Quinn, Conrad Padraig  
; APPLICANT: Foster, Keith Alan  
; APPLICANT: Chaddock, John  
; APPLICANT: Marks, Philip  
; APPLICANT: Sutton, J. Mark  
; APPLICANT: Stancombe, Patrick  
; APPLICANT: Wayne, Jonathan  
; TITLE OF INVENTION: Recombinant Toxin Fragments  
; FILE REFERENCE: 1581.0130004  
; CURRENT APPLICATION NUMBER: US/11/077,550  
; PRIOR FILING DATE: 2005-03-11  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR FILING DATE: 1996-12-27  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR FILING DATE: 1996-12-27  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR FILING DATE: 1996-08-23  
; NUMBER OF SEQ ID NOS: 179  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 130  
; LENGTH: 877  
; TYPE: PRT  
; ORGANISM: Clostridium botulinum  
US-11-077-550-130

Query Match            5.6%; Score 81; DB 7; Length 877;  
Best Local Similarity 20.4%; Pred. No. 4.4;  
Matches 68; Conservative 39; Mismatches 100; Indels 126; Gaps 16;

QY         5 RFVVGVTLSFTVFFLSDGAFSDANF-----SEGRRLGLYGISQYKVGIPNFSNFAE 55  
            :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
DB         208 RPM--LTYSNATNDVGEGRFSKFECMDPILIMHMLNHAMH--NLGYIAIPN-----D 257

QY         56 ETIPGITKKIF----ALGLDKSEINTHSNTFRSYDPTYASSF-----93  
            :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
DB         258 QTISSVTSNIFFSQYNVKLEYAEIYAFGGPTIDILPKSARKYPFEKALDYRSIAKLNS 317

QY         94 -----AGPSGIIGYY-----VNDRFVEFGSYE---NFEPERQWYPENSQYKFFALS RN 140  
            :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
DB         318 ITTANPSSFNKYIGEYKOKLRKYRFVVESSEGVTVNRNKPFVELYNELTQTTFEFNYAK- 376

QY         141 ATNSDNKFIULEN--NGVVDKSLNVNCYDIASGISPLAPYMCAGVADYIKFLGIGSLPK 198  
            :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
DB         377 IYNQVRKIYLSNVVTPVTANILDNV-YDIQN-----GFNIPK 413

QY         199 FSQVQKF-GVNY-----PLANVTMLF-----GGGYHKVVG-----228  
            :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
DB         415 SNLVLFMGQLSRNPALKRVKNPENMLYLFTFKCHKAIDGRSLYNTKDLCRELLVKNTDL 474

QY         229 -----DRHERVEIAHYHTALS 244  
            :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
DB         475 PFIGDISDVKTDFILRKDINEETEVIYYPDNV 506

RESULT 5  
US-11-077-550-126  
; Sequence 126, Application US/11077550  
; Publication No. US2005024435A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford Charles  
; APPLICANT: Quinn, Conrad Padraig  
; APPLICANT: Foster, Keith Alan  
; APPLICANT: Chaddock, John  
; APPLICANT: Marks, Philip  
; APPLICANT: Sutton, J. Mark  
; APPLICANT: Stancombe, Patrick  
; APPLICANT: Wayne, Jonathan  
; TITLE OF INVENTION: Recombinant Toxin Fragments  
; FILE REFERENCE: 1581.0130004  
; CURRENT APPLICATION NUMBER: US/11/077,550  
; PRIOR FILING DATE: 2005-03-11  
; PRIOR FILING DATE: 10/241,596  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR FILING DATE: 1996-12-27  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR FILING DATE: 1996-08-23  
; NUMBER OF SEQ ID NOS: 179  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 126  
; LENGTH: 877  
; TYPE: PRT  
; ORGANISM: clostridium botulinum  
US-11-077-550-126

Db 208 RFM--LTYSNATNDVGBGRFSKSEFCMDPILILMHNLHAMH--NLYGIAIPN-----D 257  
Qy 56 ETIPGITKKIF---ALGLDKSEINTHSNFTSRSDPTYASSF----- 93  
Db 258 QTISSTNSNIFYSQYNVKLSYAEIYAFGGPTIDILPKSARKYFEKALDYRSIAKRLNS 317  
Qy 94 -----AGFSGIIGYY-----VNDPRVFEFGSYE---NFEPERQWYPENSQSYKFFALSRLN 140  
Db 318 ITTANPSSFNKYIGEYKQKLIRKYRFVVBSSGSEVTNRNKFVELYNELTQIFTEFNAYK- 376  
Qy 141 ATNSDNKFIVLN--NGVVDKSLNVNVCYDIAGSIPPLAPYMCAGVADYIKFLGSLPK 198  
Db 377 IYNVQNRKIYLSNVYTPVTANILDDNV-YDION-----GFNIPK 414  
Qy 199 FSYQVKF-GVNY-----PLNVNTMLF-----GGGYHKVVG----- 228  
Db 415 SNLNVLFMGQNLNRNPAIRKVNPNENMLYFTFKCHKAIDGRSLYKNTLDCRELLVKNKTDL 474  
Qy 229 -----DRHERVEIAYHPTALS 244  
Db 475 PFIGDISDKTDFILRKDINEETEVYYPDNVS 507

## RESULT 7

US-11-077-550-124  
; Sequence 124, Application US/11077550  
; Publication No. US20050244435A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford Charles  
; APPLICANT: Quinn, Conrad Padraig  
; APPLICANT: Foster, Keith Alan  
; APPLICANT: Chaddock, John  
; APPLICANT: Marks, Philip  
; APPLICANT: Sutton, J. Mark  
; APPLICANT: Stancombe, Patrick  
; APPLICANT: Wayne, Jonathan  
; TITLE OF INVENTION: Recombinant Toxin Fragments  
; FILE REFERENCE: 1581.0130004  
; CURRENT APPLICATION NUMBER: US/11/077,550  
; CURRENT FILING DATE: 2005-03-11  
; PRIOR APPLICATION NUMBER: 10/241,596  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: 09/255,829  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: PCT/GB97/02273  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: 08/782,893  
; PRIOR FILING DATE: 1996-12-27  
; PRIOR APPLICATION NUMBER: GB9625996.5  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR APPLICATION NUMBER: GB9617671.4  
; PRIOR FILING DATE: 1996-08-23  
; NUMBER OF SEQ ID NOS: 179  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 124  
; LENGTH: 881  
; TYPE: PRF  
; ORGANISM: Clostridium botulinum  
US-11-077-550-124

Query Match 5.6%; Score 81; DB 7; Length 881;  
Best Local Similarity 20.4%; Pred. No. 4.4;  
Matches 68; Conservative 39; Mismatches 100; Indels 126; Gaps 16;

Qy 5 RFVGVGVTLSFTVFLSDGAFSDANF-----SEGRRLYIGSQYKVGIPNFSNFSAE 55  
Db 212 RFM--LTYSNATNDVGBGRFSKSEFCMDPILILMHNLHAMH--NLYGIAIPN-----D 261  
Qy 56 ETIPGITKKIF---ALGLDKSEINTHSNFTSRSDPTYASSF----- 93  
Db 262 QTISSTNSNIFYSQYNVKLSYAEIYAFGGPTIDILPKSARKYFEKALDYRSIAKRLNS 321  
Qy 94 -----AGFSGIIGYY-----VNDPRVFEFGSYE---NFEPERQWYPENSQSYKFFALSRLN 140

Db 322 ITTANPSSFNKYIGEYKQKLIRKYRFVVBSSGSEVTNRNKFVELYNELTQIFTEFNAYK- 380  
Qy 141 ATNSDNKFIVLN--NGVVDKSLNVNVCYDIAGSIPPLAPYMCAGVADYIKFLGSLPK 198  
Db 381 IYNVQNRKIYLSNVYTPVTANILDDNV-YDION-----GFNIPK 418  
Qy 199 FSYQVKF-GVNY-----PLNVNTMLF-----GGGYHKVVG----- 228  
Db 419 SNLNVLFMGQNLNRNPAIRKVNPNENMLYFTFKCHKAIDGRSLYKNTLDCRELLVKNKTDL 478  
Qy 229 -----DRHERVEIAYHPTALS 244  
Db 479 PFIGDISDKTDFILRKDINEETEVYYPDNVS 511

## RESULT 8

US-11-077-550-132  
; Sequence 132, Application US/11077550  
; Publication No. US20050244435A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford Charles  
; APPLICANT: Quinn, Conrad Padraig  
; APPLICANT: Foster, Keith Alan  
; APPLICANT: Chaddock, John  
; APPLICANT: Marks, Philip  
; APPLICANT: Sutton, J. Mark  
; APPLICANT: Stancombe, Patrick  
; APPLICANT: Wayne, Jonathan  
; TITLE OF INVENTION: Recombinant Toxin Fragments  
; FILE REFERENCE: 1581.0130004  
; CURRENT APPLICATION NUMBER: US/11/077,550  
; CURRENT FILING DATE: 2005-03-11  
; PRIOR APPLICATION NUMBER: 10/241,596  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: 09/255,829  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: PCT/GB97/02273  
; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: 08/782,893  
; PRIOR FILING DATE: 1996-12-27  
; PRIOR APPLICATION NUMBER: GB9625996.5  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR APPLICATION NUMBER: GB9617671.4  
; PRIOR FILING DATE: 1996-08-23  
; NUMBER OF SEQ ID NOS: 179  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 132  
; LENGTH: 902  
; TYPE: PRF  
; ORGANISM: Clostridium botulinum  
US-11-077-550-132

Query Match 5.6%; Score 81; DB 7; Length 902;  
Best Local Similarity 20.4%; Pred. No. 4.6;  
Matches 68; Conservative 39; Mismatches 100; Indels 126; Gaps 16;

Qy 5 RFVGVGVTLSFTVFLSDGAFSDANF-----SEGRRLYIGSQYKVGIPNFSNFSAE 55  
Db 212 RFM--LTYSNATNDVGBGRFSKSEFCMDPILILMHNLHAMH--NLYGIAIPN-----D 261  
Qy 56 ETIPGITKKIF---ALGLDKSEINTHSNFTSRSDPTYASSF----- 93  
Db 262 QTISSTNSNIFYSQYNVKLSYAEIYAFGGPTIDILPKSARKYFEKALDYRSIAKRLNS 321  
Qy 94 -----AGFSGIIGYY-----VNDPRVFEFGSYE---NFEPERQWYPENSQSYKFFALSRLN 140  
Db 322 ITTANPSSFNKYIGEYKQKLIRKYRFVVBSSGSEVTNRNKFVELYNELTQIFTEFNAYK- 380  
Qy 141 ATNSDNKFIVLN--NGVVDKSLNVNVCYDIAGSIPPLAPYMCAGVADYIKFLGSLPK 198  
Db 381 IYNVQNRKIYLSNVYTPVTANILDDNV-YDION-----GFNIPK 418

```
Qy 199 PSYQVKP-GVNY-----PLNVNTMLF-----GGYYHKVVG----- 228
Db 419 SNLNVLMGQNLNRNALRVKPNMMLYLFKFKCHKAIDGRSLYNTKLDCRELLVKNTDL 478
Qy 229 -----DRHERVEIAYHPTALS 244
Db 479 PFIGDISDKVTDFILRKDINEETEVIYPPDNVS 511

RESULT 9
US-11-077-550-116
; Sequence 116, Application US/11077550
; Publication No. US20050244435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT FILING DATE: 2005-03-11
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 116
; LENGTH: 912
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-11-077-550-116

Query Match 5.6%; Score 81; DB 7; Length 912;
Best Local Similarity 20.4%; Pred. No. 4.6;
Matches 68; Conservative 39; Mismatches 100; Indels 126; Gaps 16;

Qy 5 RFVVGVTLSFTVFPLSDGAFSDANF-----SEGRRLYIGSQYKVGIPNFSFSAE 55
Db 212 RFM--LTYSNATNDVGEGRFSKSEFCMDPILILMHELNHAMH--NLYGIAIPN-----D 261
Qy 56 ETIPGITKKIF-----ALGLDKSEINTHSNFTSRSDPTVASSF----- 93
Db 262 QTISSVTSNIFYSQYNVKLEAYIAFGGPTIDILPKSARKYFEEKALDYRSTAKRLNS 321
Qy 94 -----AGFSGIIGYV-----VNDPRVEFGSYE---NFEPEQWYPENSQYKPFALSRN 140
Db 322 ITTANPSSFNKYIGYKQKLIRKYRFVVESSGEVTVNRNKFVELYNELTQIFTEFNAYAK- 380
Qy 141 ATNSDNKFIULEN--NGVVDKSLNVNVCYDIASGSIPLAPYMCAGVADYIKFGLISLPK 198
Db 381 IYVQNRKIYLSNVYTPVTANILDDNV-YDIQN-----GFNIPK 418
Qy 199 FSQVQKP-GVNY-----PLNVNTMLF-----GGYYHKVVG----- 228
Db 419 SNLNVLMGQNLNRNALRVKPNMMLYLFKFKCHKAIDGRSLYNTKLDCRELLVKNTDL 478
Qy 229 -----DRHERVEIAYHPTALS 244
Db 479 PFIGDISDKVTDFILRKDINEETEVIYPPDNVS 511

RESULT 11
US-11-077-550-122
; Sequence 122, Application US/11077550
; Publication No. US20050244435A1
; GENERAL INFORMATION:
```

```
Qy 199 PSYQVKP-GVNY-----PLNVNTMLF-----GGYYHKVVG----- 228
Db 419 SNLNVLMGQNLNRNALRVKPNMMLYLFKFKCHKAIDGRSLYNTKLDCRELLVKNTDL 478
Qy 229 -----DRHERVEIAYHPTALS 244
Db 479 PFIGDISDKVTDFILRKDINEETEVIYPPDNVS 511

RESULT 9
US-11-077-550-116
; Sequence 116, Application US/11077550
; Publication No. US20050244435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT FILING DATE: 2005-03-11
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 116
; LENGTH: 912
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-11-077-550-116

Query Match 5.6%; Score 81; DB 7; Length 912;
Best Local Similarity 20.4%; Pred. No. 4.6;
Matches 68; Conservative 39; Mismatches 100; Indels 126; Gaps 16;

Qy 5 RFVVGVTLSFTVFPLSDGAFSDANF-----SEGRRLYIGSQYKVGIPNFSFSAE 55
Db 212 RFM--LTYSNATNDVGEGRFSKSEFCMDPILILMHELNHAMH--NLYGIAIPN-----D 261
Qy 56 ETIPGITKKIF-----ALGLDKSEINTHSNFTSRSDPTVASSF----- 93
Db 262 QTISSVTSNIFYSQYNVKLEAYIAFGGPTIDILPKSARKYFEEKALDYRSTAKRLNS 321
Qy 94 -----AGFSGIIGYV-----VNDPRVEFGSYE---NFEPEQWYPENSQYKPFALSRN 140
Db 322 ITTANPSSFNKYIGYKQKLIRKYRFVVESSGEVTVNRNKFVELYNELTQIFTEFNAYAK- 380
Qy 141 ATNSDNKFIULEN--NGVVDKSLNVNVCYDIASGSIPLAPYMCAGVADYIKFGLISLPK 198
Db 381 IYVQNRKIYLSNVYTPVTANILDDNV-YDIQN-----GFNIPK 418
Qy 199 FSQVQKP-GVNY-----PLNVNTMLF-----GGYYHKVVG----- 228
Db 419 SNLNVLMGQNLNRNALRVKPNMMLYLFKFKCHKAIDGRSLYNTKLDCRELLVKNTDL 478
Qy 229 -----DRHERVEIAYHPTALS 244
Db 479 PFIGDISDKVTDFILRKDINEETEVIYPPDNVS 511
```

```

; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 118
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-11-077-550-118

Query Match          5.6%; Score 81; DB 7; Length 950;
Best Local Similarity 20.4%; Pred. No. 4.9;
Matches 68; Conservative 39; Mismatches 100; Indels 126; Gaps 16;

Qy  5  RFVGVTLSTFVFLSDGAFSDANP-----SEGRRLYIGSQYKVGIPNFSNFAE 55
Db  212  RFM--LTYSNATNDVGEGRFSSEFCMDPILILMHNLHAMH--NLYGIAIPN-----D 261
Qy  56  ETIPGTTKTF-----ALGDKSEINTHSNTRSVDPTYASSF----- 93
Db  262  QTSSVTSNTIFYSQYNVVKLEYABIAYPGGPTIDLPKSARKYFEKALDYVRSIAKLNS 321
Qy  94  -----AGFSGIIGY-----VNDPRVFEFGSYE---NFEPEROWYPENSQSKPFALSRN 140
Db  322  ITTANSSFNKYIGEYKQKLIRKYRVFVSSGEVTVNRNKFVELYNELTQIFTEFNyak- 380
Qy  141  ATNSDNKFIVLN--NGVVDKLSNVNVCYDIAGSIPLAPYMCAGVADYIKFIGISLPK 198
Db  381  IYNVQRKIYLSNVYTPVTANILDDNV-YDIQN-----GFIPIK 418
Qy  199  FSQVQKF-GVNY-----PLNVNTMLF-----GGYVHKVVG----- 228
Db  419  SNLNLVPMGQNLGRNPALRKVPENMLYLFTKFKCHKAIDGRSLYNKTLDCRELLVKNITDL 478
Qy  229  -----DRERVEIAHYHTALS 244
Db  479  PFIGDISDVKTDFLRKDINEETEVIYPPDNVS 511

RESULT 13
US-11-170-653-29
; Sequence 29, Application US/11/170653
; Publication No. US2005027169A1
; GENERAL INFORMATION:
; APPLICANT: Danisco A/S
; APPLICANT: Sibbesen, Ole
; APPLICANT: Sorensen, Jens
; TITLE OF INVENTION: Xylanase Variants Having Altered Sensitivity to Xylanase
; FILE REFERENCE: 674509-2046
; CURRENT APPLICATION NUMBER: US/11/170,653
; CURRENT FILING DATE: 2005-06-23
; PRIOR APPLICATION NUMBER: US/10/237,386
; PRIOR FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: PCT/IB01/00426
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: GB 0005585.5
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: GB 0015751.1

```

```

; PRIOR FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 29
; LENGTH: 219
; TYPE: PRT
; ORGANISM: C. gracile
US-11-170-653-29

```

```

Query Match          5.4%; Score 78; DB 7; Length 219;
Best Local Similarity 20.5%; Pred. No. 1.3;
Matches 43; Conservative 32; Mismatches 75; Indels 60; Gaps 11;

Qy 9 GVTUFTVFVSDGAFDANSEGRGLYIGSQYKVGIPNFSNF-SABETIPGTTKIFPA 67
Db 38 GTNNGYFYFWTDGG-GTVNYQNG-----AGGSYSVQWQNGFVGGKWNPGGAARTI-- 89
Qy 68 LGLDKSEINTSHSTRSDPTVYASSPAGFSG-----IIGYVNDFRVEFEFGSYENFEPER 122
Db 90 -----NFSGTFSP-QGNGYLAITYGTONPLVEYXI-----VESFGTYD----- 126
Qy 123 QWYPENSQYKFFALSRLNATNSDNKFIVLNNGVVDKSLNVNVCYD-----IASGS 173
Db 127 -----PSSQASKEGTIQDQGST-----YTIKTRVNPQSIETSTFDQFWSVRQNHRSQS 178
Qy 174 IPLAPYMCAGVADYIKFLGSLPKFSYQV 203
Db 179 VNVAHFNAWAQA-----GLKLGSHNYQI 202

```

## RESULT 14

```

US-11-191-374-3
; Sequence 3, Application US/11191374
; Publication No. US20050260673A1
; GENERAL INFORMATION:
; APPLICANT: Hresko, Michelle Coutu
; APPLICANT: McLaire, Merry B.
; APPLICANT: Williams, Deryck J.
; APPLICANT: Frevert, Anita M.
; APPLICANT: Chiapelli, Brandi
; APPLICANT: Baublite, Catherine
; APPLICANT: Kloeck, Andrew P.
; APPLICANT: Davila-Aponte, Jennifer A.
; APPLICANT: Bradley, John D.
; TITLE OF INVENTION: NEMATODE PAN AND ZP RECEPTOR-LIKE
; FILE REFERENCE: 12557-015001
; CURRENT FILING DATE: 2005-07-28
; PRIOR FILING DATE: 2004-02-04
; PRIOR FILING DATE: 2004-02-04
; PRIOR FILING DATE: 2003-02-04
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1122
; TYPE: PRT
; ORGANISM: Strongyloides stercoralis
US-11-191-374-3

```

```

Query Match          5.4%; Score 77.5; DB 7; Length 1122;
Best Local Similarity 22.6%; Pred. No. 1.3;
Matches 36; Conservative 31; Mismatches 57; Indels 35; Gaps 8;

Qy 118 FEPERQWYPENSQYKFFALSRLNATNSDNKFIVLNNGVVDKSLNVNVCYDIASGSIPL 176
Db 631 YEPEKKYHPKSGKNETSYETGTVNKSNVEEVSETLTNSGVESGSLKNI--TAPPSIPK 688
Qy 177 APYMCAG-----VGADYIKFLGSLPKFSYQ-VKFGVNYPLNVTMLFGGGYHKVVG 228
Db 689 IP---EGPLPVPILIPADQVQTI-----CDYEGIKVQIKSPQSTGTGIVFNHETC-- 737

```

```

Qy 229 DRHERVEIAHY-----PTALSVDVPRTTASATLNT 258
Db 738 ----RVEVSNSDAATLELGLPASFGMKPVTLSATSSDST 772

```

## RESULT 15

```

US-11-191-375-3
; Sequence 3, Application US/11191375
; Publication No. US20050260674A1
; GENERAL INFORMATION:
; APPLICANT: Hresko, Michelle Coutu
; APPLICANT: McLaire, Merry B.
; APPLICANT: Williams, Deryck J.
; APPLICANT: Frevert, Anita M.
; APPLICANT: Chiapelli, Brandi
; APPLICANT: Baublite, Catherine
; APPLICANT: Kloeck, Andrew P.
; APPLICANT: Davila-Aponte, Jennifer A.
; APPLICANT: Bradley, John D.
; APPLICANT: Xu, Siqun
; TITLE OF INVENTION: NEMATODE PAN AND ZP RECEPTOR-LIKE
; FILE REFERENCE: 12557-015001
; CURRENT FILING DATE: 2005-07-28
; PRIOR FILING DATE: 2004-02-04
; PRIOR FILING DATE: 2004-02-04
; PRIOR FILING DATE: 2003-02-04
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1122
; TYPE: PRT
; ORGANISM: Strongyloides stercoralis
US-11-191-375-3

```

```

Query Match          5.4%; Score 77.5; DB 7; Length 1122;
Best Local Similarity 22.6%; Pred. No. 1.3;
Matches 36; Conservative 31; Mismatches 57; Indels 35; Gaps 8;

Qy 118 FEPERQWYPENSQYKFFALSRLNATNSDNKFIVLNNGVVDKSLNVNVCYDIASGSIPL 176
Db 631 YEPEKKYHPKSGKNETSYETGTVNKSNVEEVSETLTNSGVESGSLKNI--TAPPSIPK 688
Qy 177 APYMCAG-----VGADYIKFLGSLPKFSYQ-VKFGVNYPLNVTMLFGGGYHKVVG 228
Db 689 IP---EGPLPVPILIPADQVQTI-----CDYEGIKVQIKSPQSTGTGIVFNHETC-- 737
Qy 229 DRHERVEIAHY-----PTALSVDVPRTTASATLNT 258
Db 738 ----RVEVSNSDAATLELGLPASFGMKPVTLSATSSDST 772

```

```

Search completed: December 28, 2005, 15:45:23
Job time : 19 secs

```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 28, 2005, 15:38:05 ; Search time 79 Seconds  
(without alignments)

1433.313 Million cell updates/sec

Title: US-10-731-554-46

Perfect score: 1434

Sequence: 1 MNYKRFVVGVTLSFVFELS.....ASATLNTDYFGWEIGRFRAL 271

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA\_Main.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1434	100.0	271	3	US-09-811-007-46 Sequence 46, Appl
2	1434	100.0	271	4	US-10-062-624-46 Sequence 46, Appl
3	1434	100.0	271	4	US-10-062-051-46 Sequence 46, Appl
4	1434	100.0	271	4	US-10-062-920-46 Sequence 46, Appl
5	1434	100.0	271	4	US-10-680-349-46 Sequence 46, Appl
6	1434	100.0	271	5	US-10-731-554-46 Sequence 46, Appl
7	1194	83.3	271	3	US-09-846-808-20 Sequence 20, Appl
8	1194	83.3	271	4	US-10-284-986-20 Sequence 20, Appl
9	1194	83.3	271	4	US-10-369-293-20 Sequence 20, Appl
10	1194	83.3	271	4	US-10-285-042-20 Sequence 20, Appl
11	509.5	35.5	280	3	US-09-811-007-42 Sequence 42, Appl
12	509.5	35.5	280	4	US-10-062-624-42 Sequence 42, Appl
13	509.5	35.5	280	4	US-10-059-964-42 Sequence 42, Appl
14	509.5	35.5	280	4	US-10-062-051-42 Sequence 42, Appl
15	509.5	35.5	280	4	US-10-062-920-42 Sequence 42, Appl
16	509.5	35.5	280	4	US-10-314-639-48 Sequence 48, Appl
17	509.5	35.5	280	4	US-10-680-349-42 Sequence 42, Appl
18	509.5	35.5	280	5	US-10-731-554-42 Sequence 42, Appl
19	509.5	35.5	280	5	US-10-901-714-48 Sequence 48, Appl
20	509.5	35.5	280	5	US-10-846-808-14 Sequence 48, Appl
21	507	35.4	283	3	US-09-811-007-10 Sequence 10, Appl
22	507	35.4	283	4	US-10-062-624-10 Sequence 10, Appl
23	507	35.4	283	4	US-10-059-964-4 Sequence 4, Appl
24	507	35.4	283	4	US-10-062-051-10 Sequence 10, Appl
25	507	35.4	283	4	US-10-284-986-14 Sequence 14, Appl
26	507	35.4	283	4	US-10-062-920-10 Sequence 10, Appl
27	507	35.4	283	4	US-10-062-920-10 Sequence 10, Appl

28	507	35.4	283	4	US-10-314-639-4	Sequence 4, Appl
29	507	35.4	283	4	US-10-369-293-14	Sequence 14, Appl
30	507	35.4	283	4	US-10-285-042-14	Sequence 14, Appl
31	507	35.4	283	4	US-10-680-349-10	Sequence 10, Appl
32	507	35.4	283	5	US-10-731-554-10	Sequence 10, Appl
33	507	35.4	283	5	US-10-901-714-4	Sequence 4, Appl
34	507	35.4	283	5	US-10-901-774-4	Sequence 4, Appl
35	494	34.4	285	3	US-09-846-808-9	Sequence 9, Appl
36	494	34.4	285	4	US-10-284-986-9	Sequence 9, Appl
37	494	34.4	285	4	US-10-369-293-9	Sequence 9, Appl
38	494	34.4	285	4	US-10-285-042-9	Sequence 30, Appl
39	492	34.3	285	4	US-10-059-964-30	Sequence 30, Appl
40	492	34.3	285	5	US-10-314-639-30	Sequence 30, Appl
41	492	34.3	285	5	US-10-901-714-30	Sequence 30, Appl
42	492	34.3	285	5	US-10-901-774-30	Sequence 30, Appl
43	489	34.1	279	4	US-10-059-964-24	Sequence 24, Appl
44	489	34.1	279	4	US-10-314-639-24	Sequence 24, Appl
45	489	34.1	279	5	US-10-901-714-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1  
US-09-811-007-46  
; Sequence 46, Application US/09811007  
; Publication No. US20030185849A1  
; GENERAL INFORMATION:  
; APPLICANT: Walker, David H.  
; APPLICANT: McBride, Jere W.  
; APPLICANT: Yu, Xue-Jie  
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein  
; FILE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof  
; FILE REFERENCE: D6152CIP2  
; CURRENT APPLICATION NUMBER: US/09/811.007  
; CURRENT FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 09/660,587  
; PRIOR FILING DATE: 2000-09-12  
; NUMBER OF SEQ ID NOS: 46  
; SEQ ID NO 46  
; LENGTH: 271  
; TYPE: PRT  
; ORGANISM: Ehrlichia canis  
; FEATURE:  
; OTHER INFORMATION: amino acid sequence of E. canis p28-9 protein  
US-09-811-007-46

Query Match	100.0%;	Score 1434;	DB 3;	Length 271;
Best Local Similarity	100.0%;	Pred. No. 2.2e-134;		
Matches 271;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MNYKRFVVGVTLSFVFFLSDGAFSDANFSEGRRLYIGSQYKVGIPNFSNFSAEETIPG	60	
Db	1	MNYKRFVVGVTLSFVFFLSDGAFSDANFSEGRRLYIGSQYKVGIPNFSNFSAEETIPG	60	
Qy	61	ITKKIFALGDKSEINTHSNPTSYDPTASSPAGFSGLIGYVNDFRVEFGSYENPEP	120	
Db	61	ITKKIFALGDKSEINTHSNPTSYDPTASSPAGFSGLIGYVNDFRVEFGSYENPEP	120	
Qy	121	ERQWYSPNSQSKYKFFALSRLNATNSDNKFI VLENNGVVDKSLNNVVCYDIAGSGIPLAPYM	180	
Db	121	ERQWYSPNSQSKYKFFALSRLNATNSDNKFI VLENNGVVDKSLNNVVCYDIAGSGIPLAPYM	180	
Qy	181	CAGVGADYIKFLGISLPKFSYQVKFGVYVPLNVNTMLFGGGYHKVVGDRHVERVEIAYHP	240	
Db	181	CAGVGADYIKFLGISLPKFSYQVKFGVYVPLNVNTMLFGGGYHKVVGDRHVERVEIAYHP	240	
Qy	241	TALSDVPRTTSSASATLNTDYFGWEIGRFRAL	271	
Db	241	TALSDVPRTTSSASATLNTDYFGWEIGRFRAL	271	

RESULT 2

```
US-10-062-624-46
; Sequence 46, Application US/10062624
; Publication No. US20020115840A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-Kilodalton Immunodominant Protein
; FILE REFERENCE: D6152CIP2/DI
; CURRENT APPLICATION NUMBER: US/10/062,624
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 09/660,587
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 46
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: amino acid sequence of E. canis p28-9 protein
US-10-062-624-46

Query Match      100.0%; Score 1434; DB 4; Length 271;
Best Local Similarity 100.0%; Pred. No. 2.2e-134;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNYKRFVVGVTLSFVFFFLSDGAFSDANFSEGRRLGIYGSQYKVGIPNFSNFSABETIPG 60
Db 1 MNYKRFVVGVTLSFVFFFLSDGAFSDANFSEGRRLGIYGSQYKVGIPNFSNFSABETIPG 60
QY 61 ITKKIFALGLDKSEINTHSNFTSYDPTVYASSFAGSGIIGYVNDPRVFEFGSYENFEP 120
Db 61 ITKKIFALGLDKSEINTHSNFTSYDPTVYASSFAGSGIIGYVNDPRVFEFGSYENFEP 120
QY 121 ERQWYPENSQSYKFFALSRNATNSDNKFTVLENNGVVDKSLNVNVCYDIASGSIPLAPYM 180
Db 121 ERQWYPENSQSYKFFALSRNATNSDNKFTVLENNGVVDKSLNVNVCYDIASGSIPLAPYM 180
QY 181 CAGVGADYIKFLGISLPLKFSYQVKGFGVNYPLNVNTMLFGGGYYHKVVGDRHERVEIAYHP 240
Db 181 CAGVGADYIKFLGISLPLKFSYQVKGFGVNYPLNVNTMLFGGGYYHKVVGDRHERVEIAYHP 240
QY 241 TALSDVPRTTSASATLNTDYFGWEIGFRPAL 271
Db 241 TALSDVPRTTSASATLNTDYFGWEIGFRPAL 271

RESULT 4
US-10-062-920-46
; Sequence 46, Application US/10062920
; Publication No. US20030096250A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
; FILE REFERENCE: D6152CIP2
; CURRENT APPLICATION NUMBER: US/10/062,920
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US/09/660,587
; PRIOR FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 09/261,358
; PRIOR FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 46
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: amino acid sequence of E. canis p28-9 protein
US-10-062-920-46

Query Match      100.0%; Score 1434; DB 4; Length 271;
Best Local Similarity 100.0%; Pred. No. 2.2e-134;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNYKRFVVGVTLSFVFFFLSDGAFSDANFSEGRRLGIYGSQYKVGIPNFSNFSABETIPG 60
Db 1 MNYKRFVVGVTLSFVFFFLSDGAFSDANFSEGRRLGIYGSQYKVGIPNFSNFSABETIPG 60
QY 61 ITKKIFALGLDKSEINTHSNFTSYDPTVYASSFAGSGIIGYVNDPRVFEFGSYENFEP 120
Db 61 ITKKIFALGLDKSEINTHSNFTSYDPTVYASSFAGSGIIGYVNDPRVFEFGSYENFEP 120
QY 121 ERQWYPENSQSYKFFALSRNATNSDNKFTVLENNGVVDKSLNVNVCYDIASGSIPLAPYM 180
Db 121 ERQWYPENSQSYKFFALSRNATNSDNKFTVLENNGVVDKSLNVNVCYDIASGSIPLAPYM 180
QY 181 CAGVGADYIKFLGISLPLKFSYQVKGFGVNYPLNVNTMLFGGGYYHKVVGDRHERVEIAYHP 240
Db 181 CAGVGADYIKFLGISLPLKFSYQVKGFGVNYPLNVNTMLFGGGYYHKVVGDRHERVEIAYHP 240
QY 241 TALSDVPRTTSASATLNTDYFGWEIGFRPAL 271
Db 241 TALSDVPRTTSASATLNTDYFGWEIGFRPAL 271

US-10-062-051-46
; Sequence 46, Application US/10062051
; Publication No. US20030073095A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
; FILE REFERENCE: D6152CIP2
; CURRENT APPLICATION NUMBER: US/10/062,051
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US/09/660,587
; PRIOR FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 09/261,358
; PRIOR FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 46
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: amino acid sequence of E. canis p28-9 protein
US-10-062-051-46
```

```
|||||
Db      241 TALSDVPRRTTSASATLNTDYFGWEIGRFRAL 271

RESULT 5
US-10-680-349-46
; Sequence 46, Application US/10680349
; Publication No. US20040198951A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-Kilodalton Immunodominant Protein
; FILE REFERENCE: D6152CIP2/D1
; CURRENT APPLICATION NUMBER: US/10/680,349
; CURRENT FILING DATE: 2003-10-07
; PRIOR APPLICATION NUMBER: US/10/062,624
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 09/660,587
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 46
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: amino acid sequence of E. canis p28-9 protein
US-10-680-349-46

Query Match      100.0%; Score 1434; DB 4; Length 271;
Best Local Similarity 100.0%; Pred. No. 2.2e-134;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MNYKRFVVGVTLSFVFFLSDGAFSDANFSGRRGLYIGSQYKVGIPNFSNFSAEETIPG 60
Db      1 MNYKRFVVGVTLSFVFFLSDGAFSDANFSGRRGLYIGSQYKVGIPNFSNFSAEETIPG 60
Qy      61 ITKKIFALGLDKSEINTHSNFTSRSDYPTVASSFAGSGIIGYYVNDPRVFEFGSYENFEP 120
Db      61 ITKKIFALGLDKSEINTHSNFTSRSDYPTVASSFAGSGIIGYYVNDPRVFEFGSYENFEP 120
Qy      121 ERQWYPENSQSYKFFALSRNATNSDNKFIVLNNGVVDKSLNVNVCYDIASGSIPLAPYM 180
Db      121 ERQWYPENSQSYKFFALSRNATNSDNKFIVLNNGVVDKSLNVNVCYDIASGSIPLAPYM 180
Qy      181 CAGVGADYIKFLGISLKPFSYQVKGYNVPLNVTMLFGGGYHKVVGDRHERVEIAYHP 240
Db      181 CAGVGADYIKFLGISLKPFSYQVKGYNVPLNVTMLFGGGYHKVVGDRHERVEIAYHP 240
Qy      241 TALSDVPRRTTSASATLNTDYFGWEIGRFRAL 271
Db      241 TALSDVPRRTTSASATLNTDYFGWEIGRFRAL 271

RESULT 7
US-09-846-808-20
; Sequence 20, Application US/09846808
; Patent No. US20020064531A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: Yu, Xu-Jie
; TITLE OF INVENTION: Ehrlichia chaffeensis 28 kDa Outer Membrane
; FILE REFERENCE: D6311
; CURRENT APPLICATION NUMBER: US/09/846,808
; CURRENT FILING DATE: 2001-05-01
; PRIOR APPLICATION NUMBER: 60/201,035
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 20
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Ehrlichia chaffeensis
; FEATURE:
; OTHER INFORMATION: P28-20 Outer Membrane Protein of
; OTHER INFORMATION: Ehrlichia chaffeensis
US-09-846-808-20

Query Match      83.3%; Score 1194; DB 3; Length 271;
Best Local Similarity 78.6%; Pred. No. 2e-110;
Matches 213; Conservative 33; Mismatches 25; Indels 0; Gaps 0;

Qy      1 MNYKRFVVGVTLSFVFFLSDGAFSDANFSGRRGLYIGSQYKVGIPNFSNFSAEETIPG 60
Db      1 MNYKRFVVGVALATLLSFLPDNSFDANVPEGRGFYVGTQYKVGVPNFSNFSAEETLPG 60
Qy      61 ITKKIFALGLDKSEINTHSNFTSRSDYPTVASSFAGSGIIGYYVNDPRVFEFGSYENFEP 120
Db      61 LTKSIFALGLDKSISDHAGFTQAYNPTYASNFAGFGVGIIGYYVNDPRVFEFGSYENFEP 120
Qy      121 ERQWYPENSQSYKFFALSRNATNSDNKFIVLNNGVVDKSLNVNVCYDIASGSIPLAPYM 180
Db      121 ERQWYPESGESHKFFALSRSTVDNKNFIVLENDGVIDKSLNVNVCYDIAHGSIPAPYM 180
```

Qy	181	CAGVGADYIKFGLGISLPKFSYQVKEGVNPLNVNTMLFGGQYHHKVQVCDRHRERVETAYHP	240
Db	181	CAGVGADYIKFGLGISLPKFSYQVKEGVNPLNVNTMLFGGQYHHKVQVCDRHRERVETAYHP	240
Qy	241	TALSDVPRITTSASATLNTDYFCWEGIGRFPAL	271
Db	241	ATLTNNPKTTSASATLNTDYFCWEGIGRFTL	271

Query Match	83.3%;	Score 1194;	DB 4;	Length 271;
Best Local Similarity	78.6%;	Pred. No. 2e-110;		
Matches 213;	Conservative 33;	Mismatches 25;	Indels 0;	Gaps 0;
Qy	1	MNYKRFVVGVTLSFTVFPLSDGASFDANFSGRGLYIGSQYKVGIPNFSFSAETIPG	60	
Db	1	MNYKKFVVGVALATLLSPNPSFSDANVPBGRKGYVGTQYKGVPNFSFSAETIPG	60	
Qy	61	ITKKIFALGLDKSINTHSNTRSDPYTASSPAGFSGICGYVNDPRVPEGSYENPEP	120	
Db	61	LTKSIFALGLDKSISDHAGFTQYNPTYASNFAGFGVIGYVNDFRVPEGAYENPEP	120	
Qy	121	ERQWYPENSQSYKFFALSRNATNSDNKFIULENNGVDKSLNNVNCYDIASGSIPLAPYM	180	
Db	121	ERQWYPGGESHKEFALSRESVTQDNKFIULENDGVIDKSLNNVFCYDIANGSIPLAPYM	180	
Qy	181	CAGVGADYIKFLGISLPKFSQVQKFGVNNYPLNNMTLFGGGYHKVKVGDHRHVEIAYHP	240	
Db	181	CAGVGADYIKFLGISLPKFSQVQKFGVNNYPVSVMVLMFGGGYHKVIGNRYERVEIAYHP	240	
Qy	241	TALSDVPRTTASATLNTDYFGWIGRFFAL	271	
Db	241	ATLTNNPKTTASATLNTDYFGWVGMRFTL	271	

```

; SEQ ID NO 20
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Ehrlichia chaffeensis
; FEATURE:
; OTHER INFORMATION: p28-20 Outer Membrane Protein of
; OTHER INFORMATION: Ehrlichia chaffeensis
US-10-369-293-20

```

Query Match	83.3%	Score 1194;	DB 4;	Length 271;
Best Local Similarity	78.6%;	Pred. No. 2e-110;		
Matches 213;	Conservative 33;	Mismatches 25;	Indels 0;	Gaps 0;
Qy	1	MNYKRFVVGVTLTTFVFFFLSDGAFSDANSEGRGLYIGSQYKVGIPNFSNFSAEETIPG	60	
Db	1	MNYKKFVVGVALATLTLISFLPDNFSFDANVEGRKGFVQYQYKVGVPNFSNFSAEETLPG	60	
Qy	61	ITKKIIPALGJGDKSEINTHSNFTSYDPTYASSFAGPSGIIGYYVNDPRVFEGSYENFEP	120	
Db	61	LTKSIIPALGJGDKSISIDHAGFTQYNYPTYASNIFAGGGVIGYYVNDPRVFEGSYENFEP	120	
Qy	121	ERQWYPENSQSYKFFPALSRLNATNSDNKFI VLENNGVVDKSLNVCYDIAAGSGIPLAPYM	180	
Db	121	ERQWYEGGESHKFFALSRLSTVQDNKFI VLENDGVIDKSLNVCYDIAHSGIPLAPYM	180	
Qy	181	CAGVGADYIKFLGISLPKFSYQYKFGVNYPLNVTNMLFGGGYYHKVGVGDHRHERVEIAIYHP	240	
Db	181	CAGVGADYIKFLGISLPKFSYQYKFGVNYPVSVNVMFLFGGGYYHKVIGNRYERVEIAIYHP	240	
Qy	241	TALSDVPRRTTSASATLNTDYFGWEIGFRPAL	271	
Db	241	ATLTNPVKTTTSASATLTDYFGWEVGMRFTL	271	

	Query Match	83.3%	Score 1194;	DB 4;	Length 271;
	Best Local Similarity	78.6%	Pred. No. 2e-110;		
	Matches 213; Conservative	33;	Mismatches 25;	Indels 0;	Gaps 0;
Qy	1	MYNKRFFVVGVTLTFFVFFLSDGAFSDANFEGRRGLYIGSQYKVGIPNFSNFSAEETLPG	60		
Db	1	MYNKKFVVGVALATLJSLFDPNFSFDANFEGRRKFYVGTYQYKVGVPNFSNFSAEETLPG	60		
Qy	61	ITTKKIFALGDLKSEINTHSNFTSYDPTYASSPAGFSGIIGYYVNDFRVFEQSGSYENFEP	120		
Db	61	LTKSIFALGDLKSSISDHAGFTQAYNPITYASNFAFGCGVIGYYVNDFRVFEQAGAYENFEP	120		
Qy	121	ERQWYSPENSQSYKFFALSRNATNSDNKFIVLNNGVVDKSLNVNVCYDIIASGSIPLAPYIM	180		
Db	121	ERQWYPGGSGHKFFALSRSTTVODNKKFIVLENDGVIDKSLNVNFCYDIIAHSIPLAPYIM	180		

QY 181 CAGVGADYIKFLGISLPKFSYQVKGNYPLNVNTMLFGGYYHKVGDHRRHVEIAYHP 240  
DB 181 CAGVGADYIKFLGISLPKFSYQVKGNYPLNVNTMLFGGYYHKVGNRYERVEIAYHP 240  
QY 241 TALSDVPRRTTSASATLNTDYGWEIGRPFAL 271  
DB 241 ATLTVNPKRTTSASATLNTDYGWEIGRPFAL 271

## RESULT 11

US-09-811-007-42  
; Sequence 42, Application US/09811007  
; Publication No. US20030185849A1  
; GENERAL INFORMATION:  
; APPLICANT: Walker, David H.  
; APPLICANT: McBride, Jere W.  
; APPLICANT: Yu, Xue-Jie  
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein  
; FILE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof  
; FILE REFERENCE: D6152CIP2  
; CURRENT APPLICATION NUMBER: US/09/811,007  
; CURRENT FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 09/660,587  
; PRIOR FILING DATE: 2000-09-12  
; NUMBER OF SEQ ID NOS: 46  
; SEQ ID NO 42  
; LENGTH: 280  
; TYPE: PRT  
; ORGANISM: Ehrlichia canis  
; FEATURE:  
; OTHER INFORMATION: amino acid sequence of E. canis p28-2 protein  
US-09-811-007-42

Query Match 35.5%; Score 509.5; DB 3; Length 280;  
Best Local Similarity 40.5%; Pred. No. 4.1e-42;  
Matches 115; Conservative 50; Mismatches 98; Indels 21; Gaps 10;  
QY 1 MNYKRFVVGVTLSFTVFVFLSDGAFSD---ANFSEGRRLYIGSQYKVGIPNFSNFSAEET 57  
DB 1 MNYKKILVRSALISLMSILPYQSPADPVGSRNTDNKEGFYISAKYNPSISHFRKFSAEET 60  
QY 58 -IPG---ITKKIPALGLDKSEINTHNSNFTN-----SYDPTVASSFAGSGIIGYVNDFR 108  
DB 61 PINGTNSLTKKVFLGKKD-GDITKDDFTRVAPGIDFQNNLIS---GFSSGIGYSMDGPR 116  
QY 109 VEFEGSYENFEPR--QWYPENSQSYKFPALSRNATNSDNKFIVLNNGVVDKSLNVNVC 166  
DB 117 IELEAAVQOQFNPKNTDNDTNGEYKHFALSRKDMEDQQYVVLKNDGITFMSLMWNTC 176  
QY 167 YDIASGSIPLAPYMCAGVGADYIK-FLGISLPKFSYQVKGNYPLNVNTMLFGGYYHK 225  
DB 177 YDITAEVGSFVPYACAGIGADLITIFKDLNL-KFAYQKGIGISYPITPEVSFAFIGYYHG 235  
QY 226 VVGRHRRHVEIAYHTALSDVPRRTTSASATLNTDYGWEIGRPF 269  
DB 236 VIGNKFEKIPV-ITPVVLNDAPQTTSASVTLVDVGFGEIGMRP 278

## RESULT 12

US-10-062-624-42  
; Sequence 42, Application US/10062624  
; Publication No. US20020115840A1  
; GENERAL INFORMATION:  
; APPLICANT: Walker, David H.  
; APPLICANT: McBride, Jere W.  
; APPLICANT: Yu, Xue-Jie  
; TITLE OF INVENTION: Homologous 28-Kilodalton Immunodominant Protein  
; FILE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof  
; FILE REFERENCE: D6152CIP2/D1  
; CURRENT APPLICATION NUMBER: US/10/062,624  
; CURRENT FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: 09/660,587

; PRIOR FILING DATE: 2000-09-12  
; NUMBER OF SEQ ID NOS: 46  
; SEQ ID NO 42  
; LENGTH: 280  
; TYPE: PRT  
; ORGANISM: Ehrlichia canis  
; FEATURE:  
; OTHER INFORMATION: amino acid sequence of E. canis p28-2 protein  
US-10-062-624-42

Query Match 35.5%; Score 509.5; DB 4; Length 280;  
Best Local Similarity 40.5%; Pred. No. 4.1e-42;  
Matches 115; Conservative 50; Mismatches 98; Indels 21; Gaps 10;  
QY 1 MNYKRFVVGVTLSFTVFVFLSDGAFSD---ANFSEGRRLYIGSQYKVGIPNFSNFSAEET 57  
DB 1 MNYKKILVRSALISLMSILPYQSPADPVGSRNTDNKEGFYISAKYNPSISHFRKFSAEET 60  
QY 58 -IPG---ITKKIPALGLDKSEINTHNSNFTN-----SYDPTVASSFAGSGIIGYVNDFR 108  
DB 61 PINGTNSLTKKVFLGKKD-GDITKDDFTRVAPGIDFQNNLIS---GFSSGIGYSMDGPR 116  
QY 109 VEFEGSYENFEPR--QWYPENSQSYKFPALSRNATNSDNKFIVLNNGVVDKSLNVNVC 166  
DB 117 IELEAAVQOQFNPKNTDNDTNGEYKHFALSRKDMEDQQYVVLKNDGITFMSLMWNTC 176  
QY 167 YDIASGSIPLAPYMCAGVGADYIK-FLGISLPKFSYQVKGNYPLNVNTMLFGGYYHK 225  
DB 177 YDITAEVGSFVPYACAGIGADLITIFKDLNL-KFAYQKGIGISYPITPEVSFAFIGYYHG 235  
QY 226 VVGRHRRHVEIAYHTALSDVPRRTTSASATLNTDYGWEIGRPF 269  
DB 236 VIGNKFEKIPV-ITPVVLNDAPQTTSASVTLVDVGFGEIGMRP 278

## RESULT 13

US-10-059-964-48  
; Sequence 48, Application US/10059964  
; Publication No. US20020120115A1  
; GENERAL INFORMATION:  
; APPLICANT: Rikihisa, Yasuko  
; APPLICANT: Ohashi, No. US20020120115A10  
; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia  
; FILE OF INVENTION: Chafteensis  
; FILE REFERENCE: 22727/04021  
; CURRENT APPLICATION NUMBER: US/10/059,964  
; CURRENT FILING DATE: 2002-01-28  
; EARLIER APPLICATION NUMBER: 09/314,701  
; EARLIER FILING DATE: 1999-05-19  
; NUMBER OF SEQ ID NOS: 66  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 48  
; LENGTH: 280  
; TYPE: PRT  
; ORGANISM: Ehrlichia canis  
US-10-059-964-48

Query Match 35.5%; Score 509.5; DB 4; Length 280;  
Best Local Similarity 40.5%; Pred. No. 4.1e-42;  
Matches 115; Conservative 50; Mismatches 98; Indels 21; Gaps 10;

QY 1 MNYKRFVVGVTLSFTVFVFLSDGAFSD---ANFSEGRRLYIGSQYKVGIPNFSNFSAEET 57  
DB 1 MNYKKILVRSALISLMSILPYQSPADPVGSRNTDNKEGFYISAKYNPSISHFRKFSAEET 60  
QY 58 -IPG---ITKKIPALGLDKSEINTHNSNFTN-----SYDPTVASSFAGSGIIGYVNDFR 108  
DB 61 PINGTNSLTKKVFLGKKD-GDITKDDFTRVAPGIDFQNNLIS---GFSSGIGYSMDGPR 116  
QY 109 VEFEGSYENFEPR--QWYPENSQSYKFPALSRNATNSDNKFIVLNNGVVDKSLNVNVC 166  
DB 117 IELEAAVQOQFNPKNTDNDTNGEYKHFALSRKDMEDQQYVVLKNDGITFMSLMWNTC 176



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model  
Run on: December 28, 2005, 13:52:57 ; Search time 90 Seconds  
(without alignments)  
1323.018 Million cell updates/sec

Title: US-10-731-554-46  
Perfect score: 1434  
Sequence: 1 MNVRFVGVGTLSTFVFLS.....ASATLNTDYFGWIGRFRPAL 271

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 2443163 seqs, 439378781 residues  
Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_21.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*  
9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1434	100.0	271	5 AAU96118	Aau96118 Ehrlichia
2	1194	83.3	271	5 AAU73419	Aau73419 Ehrlichia
3	509.5	35.5	280	5 AAU96116	Aau96116 Ehrlichia
4	509.5	35.5	280	5 ABG77958	Abg77958 Ehrlichia
5	509.5	35.5	280	6 ADA09781	Ada09781 E. canis
6	509.5	35.5	280	9 ADW04274	Adw04274 Ehrlichia
7	507	35.4	283	5 AAU06944	Aay06944 E. chaffe
8	507	35.4	283	5 AAU96106	Aau96106 Ehrlichia
9	507	35.4	283	5 AAU73413	Aau73413 Ehrlichia
10	507	35.4	283	5 ABG77936	Abg77936 Ehrlichia
11	507	35.4	283	6 ADA09737	Ada09737 E. chaffe
12	507	35.4	283	9 ADW04230	Adw04230 Ehrlichia
13	494	34.4	285	2 AAY06957	Aay06957 E. chaffe
14	494	34.4	285	5 AAU73408	Aau73408 Ehrlichia
15	494	34.4	285	5 ABG77949	Abg77949 Ehrlichia
16	492	34.3	285	6 ADA09763	Ada09763 E. chaffe
17	492	34.3	285	9 ADW04256	Adw04256 Ehrlichia
18	489	34.1	279	2 AAY06954	Aay06954 E. chaffe
19	489	34.1	279	5 ABG77946	Abg77946 Ehrlichia
20	489	34.1	279	6 ADA09757	Ada09757 E. chaffe
21	489	34.1	279	9 ADW04250	Adw04250 Ehrlichia
22	482	33.6	279	5 AAU73405	Aau73405 Ehrlichia
23	463	32.3	283	2 AAY06955	Aay06955 E. chaffe
24	463	32.3	283	5 AAU73406	Aau73406 Ehrlichia

25	463	32.3	283	6 ADA09759	Ada09759 E. chaffe
26	463	32.3	283	9 ADW04252	Adw04252 Ehrlichia
27	455.5	31.8	298	5 AAU73410	Aau73410 Ehrlichia
28	455.5	31.8	298	5 ABG77960	Abg77960 Ehrlichia
29	455.5	31.8	298	6 ADA09785	Ada09785 E. chaffe
30	455.5	31.8	298	9 ADW04278	Adw04278 Ehrlichia
31	455	31.7	281	5 ABG77963	Abg77963 Ehrlichia
32	455	31.7	281	6 ADA09791	Ada09791 E. canis
33	455	31.7	281	9 ADW04284	Adw04284 Ehrlichia
34	455	31.7	283	5 ABG77947	Abg77947 Ehrlichia
35	447.5	31.2	300	5 AAU73411	Aau73411 Ehrlichia
36	440	30.7	291	5 ABG77943	Abg77943 Ehrlichia
37	440	30.7	291	6 ADA09751	Ada09751 E. chaffe
38	440	30.7	291	9 ADW04244	Adw04244 Ehrlichia
39	437	30.5	293	2 AAY06965	Aay06965 E. canis
40	437	30.5	293	5 AAU96115	Aau96115 Ehrlichia
41	437	30.5	293	5 ABG77956	Abg77956 Ehrlichia
42	435.5	30.4	288	2 AAY06959	Aay06959 E. canis
43	435.5	30.4	288	5 ABG77950	Abg77950 Ehrlichia
44	435.5	30.4	288	6 ADA09765	Ada09765 E. canis
45	435.5	30.4	288	9 ADW04258	Adw04258 Ehrlichia

ALIGNMENTS

RESULT 1  
AAU96118  
ID AAU96118 standard; protein; 271 AA.  
XX  
AC AAU96118;  
XX  
DT 02-JUL-2002 (first entry)  
XX  
DE Ehrlichia canis p28-9.  
XX  
KW Ehrlichia canis infection; vaccine; serodiagnostic; p28; antibacterial.  
XX  
OS Ehrlichia canis.  
XX  
PN WO200222782-A2.  
XX  
PD 21-MAR-2002.  
XX  
PF 12-SEP-2001; 2001WO-US028759.  
XX  
PR 12-SEP-2000; 2000US-00660587.  
XX  
PA (RERE-) RES DEV FOUND.  
XX  
PI Walker DH, Yu X, McBride JW;  
XX  
DR WPI; 2002-351882/38.  
XX  
N-PSDB; ABK68878.  
XX  
PT New recombinant homologous 28 kilodalton immunodominant protein from  
XX  
PT Ehrlichia canis, useful for treating Ehrlichia canis infections.  
XX  
PS Claim 16; Fig 16; 106pp; English.  
XX  
CC The invention relates to a recombinant homologous 28 kDa immunodominant  
XX  
CC protein, p28, (I), of Ehrlichia canis. (II), a 28-kDa antigen preferably  
XX  
CC dispersed in a pharmaceutically acceptable carrier, is useful for  
XX  
CC inhibiting E. canis infection in a subject. (I) is useful in the  
XX  
CC development of vaccines and serodiagnosis that are particularly  
XX  
CC effective for disease prevention and serodiagnosis. AAU96100-AAU96118  
XX  
CC represent the 28-kDa antigen amino acid sequences of the invention  
XX  
SQ Sequence 271 AA;

Query Match 100.0%; Score 1434; DB 5; Length 271;  
Best Local Similarity 100.0%; Pred. No 9.3e-145;  
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNYKRFVGVGTLSTFVFFFLSDGAFSDANFSEGRGLYIGSOYKVGIPNFSNFSABETIPG 60  
 DB 1 MNYKRFVGVGTLSTFVFFFLSDGAFSDANFSEGRGLYIGSOYKVGIPNFSNFSABETIPG 60  
 QY 61 ITKKIFALGLDKSEINTHSNFTSYDPTVASSPAGFSGIIGYVNDPFRVEFGSYENFEP 120  
 DB 61 ITKKIFALGLDKSEINTHSNFTSYDPTVASSPAGFSGIIGYVNDPFRVEFGSYENFEP 120  
 QY 121 EROWYPENSQSYKFFALSRNATNSDNKFIVLNNGVWDKSLNVCYDIASGSIPLAPYM 180  
 DB 121 EROWYPENSQSYKFFALSRNATNSDNKFIVLNNGVWDKSLNVCYDIASGSIPLAPYM 180  
 QY 181 CAGVGADYIKFLGISLPKFSYQVKGVPYPLNVTMLFGSGYYHKVVGDRHERVEIAYHP 240  
 DB 181 CAGVGADYIKFLGISLPKFSYQVKGVPYPLNVTMLFGSGYYHKVVGDRHERVEIAYHP 240  
 QY 241 TALSDVPRTTSASATLNTDYFGWEIGFRPAL 271  
 DB 241 TALSDVPRTTSASATLNTDYFGWEIGFRPAL 271

RESULT 2  
 AAU73419  
 ID AAU73419 standard; protein; 271 AA.  
 XX  
 AC AAU73419;  
 XX  
 DT 12-MAR-2002 (first entry)  
 XX  
 DE Ehrlichia chaffeensis outer membrane protein P28-20.  
 XX  
 KW Ehrlichia; outer membrane protein; P28; antibiotic; vaccine.  
 XX  
 OS Ehrlichia chaffeensis.  
 XX  
 PN WO200183699-A2.  
 XX  
 PD 08-NOV-2001.  
 XX  
 PF 01-MAY-2001; 2001WO-US013997.  
 XX  
 PR 01-MAY-2000; 2000US-0201035P.  
 XX  
 PA (RERE-) RES DEV FOUND.  
 XX  
 PI Walker DH, Yu X;  
 XX  
 DR WPI; 2002-066527/09.  
 XX  
 PT Novel Ehrlichia chaffeensis 28-kDa outer membrane protein, designated P28  
 PT useful as a vaccine against Ehrlichia chaffeensis.  
 XX  
 PS Claim 10; Fig 2; 97pp; English.  
 XX  
 CC The invention relates to isolated and purified 28-kDa outer membrane  
 CC proteins (P28-1 to P28-21) of Ehrlichia chaffeensis. P28 proteins are  
 CC encoded by a 28kDa outer membrane protein multigene family. P28 proteins  
 CC are useful as a vaccine against E.chaffeensis. DNA encoding P28 is useful  
 CC for transfecting a host cell. AAU73400-AAU73420 represent Ehrlichia  
 CC chaffeensis P28 outer membrane proteins of the invention  
 XX  
 SQ Sequence 271 AA;

Query Match 83.3%; Score 1194; DB 5; Length 271;  
 Best Local Similarity 78.6%; Pred. No. 5.2e-119;  
 Matches 213; Conservative 33; Mismatches 25; Indels 0; Gaps 0;  
 QY 1 MNYKRFVGVGTLSTFVFFFLSDGAFSDANFSEGRGLYIGSOYKVGIPNFSNFSABETIPG 60  
 DB 1 MNYKRFVGVGTLSTFVFFFLSDGAFSDANFSEGRGLYIGSOYKVGIPNFSNFSABETIPG 60  
 QY 61 ITKKIFALGLDKSEINTHSNFTSYDPTVASSPAGFSGIIGYVNDPFRVEFGSYENFEP 120

DB 61 LTKSIFALGLDKSSISDHAGFTQAVNPTVASNFAGGVIIGYVNDPFRVEFGAYENFEP 120  
 QY 121 EROWYPENSQSYKFFALSRNATNSDNKFIVLNNGVWDKSLNVCYDIASGSIPLAPYM 180  
 DB 121 EROWYPEGESHKFFALSRNATNSDNKFIVLNNGVWDKSLNVCYDIASGSIPLAPYM 180  
 QY 181 CAGVGADYIKFLGISLPKFSYQVKGVPYPLNVTMLFGSGYYHKVVGDRHERVEIAYHP 240  
 DB 181 CAGVGADYIKFLGISLPKFSYQVKGVPYPLNVTMLFGSGYYHKVVGDRHERVEIAYHP 240  
 QY 241 TALSDVPRTTSASATLNTDYFGWEIGFRPAL 271  
 DB 241 ATLTVPKTTSASATLNTDYFGWEIGFRPAL 271

RESULT 3  
 AAU96116  
 ID AAU96116 standard; protein; 280 AA.  
 XX  
 AC AAU96116;  
 XX  
 DT 02-JUL-2002 (first entry)  
 XX  
 DE Ehrlichia canis p28-2.  
 XX  
 KW Ehrlichia canis infection; vaccine; serodiagnostic; p28; antibacterial.  
 XX  
 OS Ehrlichia canis.  
 XX  
 PN WO200222782-A2.  
 XX  
 PD 21-MAR-2002.  
 XX  
 PF 12-SEP-2001; 2001WO-US028759.  
 XX  
 PR 12-SEP-2000; 2000US-00660587.  
 XX  
 PA (RERE-) RES DEV FOUND.  
 XX  
 PI Walker DH, Yu X, McBride JW;  
 XX  
 DR WPI; 2002-351882/38.  
 DR N-PSDB; ABK68876.  
 XX  
 PT New recombinant homologous 28 kilodalton immunodominant protein from  
 PT Ehrlichia canis, useful for treating Ehrlichia canis infections.  
 XX  
 PS Claim 16; Fig 14; 106pp; English.  
 XX  
 CC The invention relates to a recombinant homologous 28 kDa immunodominant  
 CC protein, P28, (I), of Ehrlichia canis. (I), a 28-kDa antigen preferably  
 CC dispersed in a pharmaceutically acceptable carrier, is useful for  
 CC inhibiting E. canis infection in a subject. (I) is useful in the  
 CC development of vaccines and serodiagnostics that are particularly  
 CC effective for disease prevention and serodiagnosis. AAU96100-AAU96118  
 CC represent the 28-kDa antigen amino acid sequences of the invention  
 XX  
 SQ Sequence 280 AA;

Query Match 35.5%; Score 509.5; DB 5; Length 280;  
 Best Local Similarity 40.5%; Pred. No. 1.4e-45;  
 Matches 115; Conservative 50; Mismatches 98; Indels 21; Gaps 10;  
 QY 1 MNYKRFVGVGTLSTFVFFFLSDGAFSDANFSEGRGLYIGSOYKVGIPNFSNFSABET 57  
 DB 1 MNYKILVRSALISLMSILPQSPADPVGSRDNDNKEGYISAKYNPSISHPKFSABET 60  
 QY 58 -IPG---ITKKIFALGLDKSEINTHSNFT---SYDPTVASSPAGFSGIIGYVNDPFR 108  
 DB 61 PINGTNSLTAKVFGUKKO-GDITKKDDPFRVAPGIDFQNNLIS---GFSGSGIYSMDGPR 116  
 QY 109 VEFEGSYENFEPER--QWYPENSQSYKFFALSRNATNSDNKFIVLNNGVWDKSLNVCY 166



Db 117 IELEAAAYQQFNPKNNTDNDNGEYKHFALSRKDMEDQQYVVLKNDGITFMSLWNTC 176  
 Qy 167 YDIASGSIPLAPYMCAGVADYIK-FLGISLPKFSYQVKGVNYPLNVNTMLFGGGYYHK 225  
 Db 177 YDITAGVSPVPYACAGIGADLITIFKDLNL-KFAYQKGIGISYPIITPEVSFAFIGGYHG 235  
 Qy 226 VVGDHREVEIAHPTALSDVPRTTSSATLNTDYFCWEIGFRF 269  
 Db 236 VIGNKFEKIPV-ITPVVLNDAPQTTASVTLVDVYFGGEGIMRF 278

## RESULT 4

ABG77958  
 ID ABG77958 standard; protein; 280 AA.

XX  
 AC ABG77958;

XX 15-NOV-2002 (first entry)

XX Ehrlichia canis outer membrane protein (P30F) #9.

XX Outer membrane protein; OMP; P30F; ehrlichiosis; infection.

XX Ehrlichia canis.

XX US2002120115-A1.

XX 29-AUG-2002.

XX 28-JAN-2002; 2002US-00059964.

XX 19-MAY-1999; 99US-00314701.

XX (RIKI/) RIKIHISA Y.

XX (ORAS/) OHASHI N.

XX Rikihisa Y, Ohashi N;

XX WPI; 2002-618954/66.

XX N-PSDB; ABS63299.

XX Isolated polynucleotide encoding an outer membrane protein of E.canis or

XX E.chaffeensis used in the diagnosis of infection.

XX Claim 10; Fig 30B; 49pp; English.

XX The invention relates to an isolated polynucleotide encoding an outer  
 CC membrane protein (OMP) of Ehrlichia canis or Ehrlichia chaffeensis used  
 CC in the diagnosis of infection. An infection such as human ehrlichiosis or  
 CC canine ehrlichiosis can be diagnosed by providing a serum sample from the  
 CC patient, providing a polypeptide or mixture of polypeptides, contacting  
 CC the sample with the polypeptide and assaying for the formation of a  
 CC complex between antibodies in the serum sample and the polypeptide, where  
 CC formation of a complex is indicative of infection with E. chaffeensis.  
 CC This sequence represents an Ehrlichia outer membrane protein of the  
 CC invention

XX Sequence 280 AA;

Query Match 35.5%; Score 509.5; DB 5; Length 280;  
 Best Local Similarity 40.5%; Pred. No. 1.4e-45;  
 Matches 115; Conservative 50; Mismatches 98; Indels 21; Gaps 10;

Qy 1 MNYKRFVVGVTLSITVFFFLSDGAFSD---ANFSEGRGLYIGSQYKVGIPNFSFSAET 57

Db 1 MNYKKILVRSALISLSILPYQSFADPVGSRNTDNKGFYISAKYNPSISHFRKFSABET 60

Qy 58 -IPG---ITKKIFALGLDKSEINTHSNFR-----SYDPTVASSPAGFSGIIGYVNDPR 108

Db 61 PINGTNSITKKVFGGLKGD-GDITKDDTRVAPGIDFQNNLIS---GFGSGISGYMDGPR 116

Qy 109 VEFEGSYENFEFER--QWYPENSQSYKFFALSRNATNSDNKPIVLNNGVVDKSLNVNC 166

Db 117 IELEAAAYQQFNPKNNTDNDNGEYKHFALSRKDMEDQQYVVLKNDGITFMSLWNTC 176  
 Qy 167 YDIASGSIPLAPYMCAGVADYIK-FLGISLPKFSYQVKGVNYPLNVNTMLFGGGYYHK 225  
 Db 177 YDITAGVSPVPYACAGIGADLITIFKDLNL-KFAYQKGIGISYPIITPEVSFAFIGGYHG 235  
 Qy 226 VVGDHREVEIAHPTALSDVPRTTSSATLNTDYFCWEIGFRF 269  
 Db 236 VIGNKFEKIPV-ITPVVLNDAPQTTASVTLVDVYFGGEGIMRF 278

## RESULT 5

ADA09781  
 ID ADA09781 standard; protein; 280 AA.

XX  
 AC ADA09781;

XX 06-NOV-2003 (first entry)

XX E. canis outer membrane protein P30-10.

XX outer membrane protein; circulating leukocyte; monocytic ehrlichiosis;  
 XX Rocky Mountain spotted fever; canine ehrlichiosis; antigen.

XX Ehrlichia canis.

XX US65444517-B1.

XX 08-APR-2003.

XX 19-MAY-1999; 99US-00314701.

XX 18-SEP-1998; 98US-0100843P.

XX (OHIS ) UNIV OHIO STATE RES FOUND.

XX Rikihisa Y, Ohashi N;

XX WPI; 2003-553952/52.

XX N-PSDB; ADA09780.

XX New isolated polynucleotide encoding outer membrane protein P30 of  
 PT Ehrlichia canis or its variant or fragment, useful for producing  
 PT Ehrlichia canis or Ehrlichia chaffeensis outer membrane proteins for  
 PT diagnosing and treating ehrlichiosis.

XX Disclosure; Fig 30; 105pp; English.

XX The invention relates to an isolated polynucleotide encoding a variant of  
 CC the outer membrane protein (OMP) P30 of Ehrlichia canis, an outer  
 CC membrane protein of E. canis , or an antigenic fragment of the E. canis  
 CC P30 protein, or comprising a sequence which is the complement of  
 CC nucleotide 99 through nucleotide 672 of ADA09764. Also disclosed are E.  
 CC canis and E. chaffeensis outer membrane proteins and their encoding  
 CC nucleic acids. The polynucleotides are useful for producing E. canis or  
 CC E. chaffeensis outer membrane protein, for designing hybridisation probes  
 CC for isolating and identifying cDNA and genomic clones encoding the OMP)  
 CC or its allelic forms, for designing primers for PCR. The polypeptides  
 CC encoded by the polynucleotide is useful for diagnosing human ehrlichiosis  
 CC (Rocky mountain spotted fever) or canine ehrlichiosis. The present  
 CC sequence represents an E. canis outer membrane protein.

XX Sequence 280 AA;

Query Match 35.5%; Score 509.5; DB 6; Length 280;

Best Local Similarity 40.5%; Pred. No. 1.4e-45;

Matches 115; Conservative 50; Mismatches 98; Indels 21; Gaps 10;

Qy 1 MNYKRFVVGVTLSITVFFFLSDGAFSD---ANFSEGRGLYIGSQYKVGIPNFSFSAET 57

Db 1 MNYKKILVRSALISLSILPYQSFADPVGSRNTDNKGFYISAKYNPSISHFRKFSABET 60





```

Db      119  RIELEAAAYQKFDKN---PDNDNTNSGDYKYFGLSREDAIADKKVVLKNEGITPMSELM 175
QY      163  VNVCYDIASSIPLAPYMCAGVADYIK-FLGISLPKFSYQVKGWYPLNVNTMLFGGG 221
Db      176  VNTCYDITABGVFFIPYACAGVADLINVKDFNL-KFSYQKGIGISYPTITPEVSFAFIG 234
QY      222  YHKWVGDRHERVEIAVHPALTSDVPRTTSSASATLNTDYFGWGEIGRPF 269
Db      235  YHGVIGNNFKIPV-ITPVLEGAQTTSALTVIDTGTGFGVGVGRF 281

```

## RESULT 10

ABG77936

ID ABG77936 standard; protein; 283 AA.

XX

AC ABG77936;

XX

DT 15-NOV-2002 (first entry)

XX

DE Ehrlichia chaffeensis outer membrane protein (OMP) #2.

XX

KW Outer membrane protein; OMP; P30F; ehrlichiosis; infection.

XX

OS Ehrlichia chaffeensis.

XX

PN US2002120115-A1.

XX

PD 29-AUG-2002.

XX

XX 28-JAN-2002; 2002US-00059964.

XX

PF 19-MAY-1999; 99US-00314701.

XX

XX (RIKI/) RIKIHISA Y.

PA (OHAS/) OHASHI N.

XX

XX Rikihisa Y, Ohashi N;

XX

DR WPI; 2002-618954/66.

XX

DR N-PSDB; ABS63277.

XX

XX Isolated polynucleotide encoding an outer membrane protein of E.canis or

PT E.chaffeensis used in the diagnosis of infection.

PT

PS Disclosure; Fig 4B; 49pp; English.

XX

CC The invention relates to an isolated polynucleotide encoding an outer  
 CC membrane protein (OMP) of Ehrlichia canis or Ehrlichia chaffeensis used  
 CC in the diagnosis of infection. An infection such as human ehrlichiosis or  
 CC canine ehrlichiosis can be diagnosed by providing a serum sample from the  
 CC patient, providing a polypeptide or mixture of polypeptides, contacting  
 CC the sample with the polypeptide and assaying for the formation of a  
 CC complex between antibodies in the serum sample and the polypeptide, where  
 CC formation of a complex is indicative of infection with E. chaffeensis.  
 CC This sequence represents an Ehrlichia outer membrane protein of the  
 CC invention

XX

SQ Sequence 283 AA;

XX

Query Match 35.4%; Score 507; DB 5; Length 283;

Best Local Similarity 40.3%; Pred. No. 2.7e-45;

Matches 116; Conservative 46; Mismatches 100; Indels 26; Gaps 11;

XX

QY 1 MNYKRFVVGVTLTSTVFVFLSDGAF-----SDANFSEGRGLYIGSQYKVGIPNFSNFA 54

Db 1 MNYKKIFVSSALISLMSILPYQSFADPVTNSDGTGINDSREGFYISVKYNPSISHFRKFA 60

XX

QY 55 EET-IPG---ITKKIFALGDKSEINTHSNFTSYDPTVA---SSPAGFGSIIGYVNDVF 107

Db 61 EEAPINGNTSITKKVFLGKKD-GDIAQSANFNT-DPALEFQNNLISGFSGSLGYANDGP 118

XX

QY 108 RVEFEGSYENFEPERQWYPENSQS-----YKFFALSRNATNSDNKFTVLNNGVVDKSLN 162

XX

```

Db      119  RIELEAAAYQKFDKN---PDNDNTNSGDYKYFGLSREDAIADKKVVLKNEGITPMSELM 175
QY      163  VNVCYDIASSIPLAPYMCAGVADYIK-FLGISLPKFSYQVKGWYPLNVNTMLFGGG 221
Db      176  VNTCYDITABGVFFIPYACAGVADLINVKDFNL-KFSYQKGIGISYPTITPEVSFAFIG 234
QY      222  YHKWVGDRHERVEIAVHPALTSDVPRTTSSASATLNTDYFGWGEIGRPF 269
Db      235  YHGVIGNNFKIPV-ITPVLEGAQTTSALTVIDTGTGFGVGVGRF 281

```

## RESULT 11

ADA09737

ID ADA09737 standard; protein; 283 AA.

XX

AC ADA09737;

XX

DT 06-NOV-2003 (first entry)

XX

DE E. chaffeensis outer membrane protein OMP-1B.

XX

KW outer membrane protein; circulating leukocyte; monocytic ehrlichiosis;

XX Rocky Mountain spotted fever; canine ehrlichiosis; antigen.

XX

OS Ehrlichia chaffeensis.

XX

PN US6544517-B1.

XX

PD 08-APR-2003.

XX

XX 19-MAY-1999; 99US-00314701.

XX

XX 18-SEP-1998; 98US-0100843P.

XX

XX (OHIS ) UNIV OHIO STATE RES FOUND.

XX

XX Rikihisa Y, Ohashi N;

XX

DR WPI; 2003-553952/52.

XX

DR N-PSDB; ADA09736.

XX

PT New isolated polynucleotide encoding outer membrane protein P30 of  
 PT Ehrlichia canis or its variant or fragment, useful for producing  
 PT Ehrlichia canis or Ehrlichia chaffeensis outer membrane proteins for  
 PT diagnosing and treating ehrlichiosis.

XX

PS Disclosure; Fig 4; 105pp; English.

XX

CC The invention relates to an isolated polynucleotide encoding a variant of  
 CC the outer membrane protein (OMP) P30 of Ehrlichia canis, an outer  
 CC membrane protein of E. canis , or an antigenic fragment of the E. canis  
 CC P30 protein, or comprising a sequence which is the complement of  
 CC nucleotide 99 through nucleotide 672 of ADA09764. Also disclosed are E.  
 CC canis and E. chaffeensis outer membrane proteins and their encoding  
 CC nucleic acids. The polynucleotides are useful for producing E. canis or  
 CC E. chaffeensis outer membrane protein, for designing hybridisation probes  
 CC for isolating and identifying cDNA and genomic clones encoding the OMP)  
 CC or its allelic forms, for designing primers for PCR. The polypeptides  
 CC encoded by the polynucleotide is useful for diagnosing human ehrlichiosis  
 CC (Rocky mountain spotted fever) or canine ehrlichiosis. The present  
 CC sequence represents an E. chaffeensis outer membrane protein.

XX

SQ Sequence 283 AA;

XX

Query Match 35.4%; Score 507; DB 6; Length 283;

Best Local Similarity 40.3%; Pred. No. 2.7e-45;

Matches 116; Conservative 46; Mismatches 100; Indels 26; Gaps 11;

XX

QY 1 MNYKRFVVGVTLTSTVFVFLSDGAF-----SDANFSEGRGLYIGSQYKVGIPNFSNFA 54

Db 1 MNYKKIFVSSALISLMSILPYQSFADPVTNSDGTGINDSREGFYISVKYNPSISHFRKFA 60

XX

Qy	55	EET-IPG---ITKKIFALGLDKSEINTHSNTRSYDPTYA---SSPAGPGGIIGYVNDP	107
Db	61	EEAPINGNTSITKKVFGGLKKD-GDIAQSANFNRT-PPALBFQNNLISFGSGSIGYAMDGP	118
Qy	108	RVEFGSYENPFPERQWTPENSQS-----YKFPALSRNATNSDNKFTVLNENGVVDKSLN	162
Db	119	RIELEAAYQKPAKQ---PDNNDTNSGDYKYGFLSRREDATADKKYVVLNKGEGITPMSLM	175
Qy	163	VNVCYDIASGSIPLAPYMCAGVGADYIK-FLGISLPKFSYQKGVNYPNLNVTMLFGGG	221
Db	176	VNTCYDITAEGVFPPIYACAGVGADLIINVKDFNL-KFSYQCKIGISYPTPEVSAFIGG	234
Qy	222	YHKVVGDRHREVEIAYHPTLASDVPRTTSSATLNTDYFGWEIGPFRF	269
Db	235	YHGVGIIGNPFNKIPV-IIPVVLFGAPOTTSSALVTIDTGYFGCEYGVGF	281

RESULT 12	
ADW04230	
ID	ADW04230 standard; protein; 283 AA.
XX	
AC	ADW04230;
XX	
XX	
DT	24-MAR-2005 (first entry)
XX	
DE	Erlichia chaffeensis outer membrane protein (OMP), OMP-1B.
XX	
XX	DNA purification; diagnosis; outer membrane protein; OMP; P30F protein;
KW	infection; vaccine.
KW	
XX	
OS	Ehrlichia chaffeensis.
XX	
Key	Location/Qualifiers
FH	
Peptide	1. .25
FT	/label= Signal_peptide
FT	
FT	Protein
FT	26. .283
FT	/note= "Erlichia chaffeensis mature OMP protein"
FT	
FT	Region
FT	26. .41
FT	/note= "Semivariable region (SV) "
FT	
FT	Region
FT	82. .94
FT	/note= "Hypervariable region (HV1) "
FT	
FT	Region
FT	145. .163
FT	/note= "Hypervariable region (HV2) "
FT	
FT	Region
FT	248. .272
FT	/note= "Hypervariable region (HV3) "
FT	

PN	US2004265334-A1.	
XX		
XX		
PD	30-DEC-2004.	
XX		
XX		
PF	29-JUL-2004; 2004US-00901774.	
XX		
XX		
PR	18-SEP-1998; 98US-0100843P.	
PR	19-MAY-1999; 99US-00314701.	
PR	28-JAN-2002; 2002US-00059964.	
XX		
PA	(RIKI/) RIKIHISA Y.	
PA	(CHAS/) OHASHI N.	
XX		
PI	Rikihisa Y, Ohashi N;	
XX		
XX	WPI; 2005-064871/07.	
DR	N-PSDB; ADW04229.	
XX		
PT	New polynucleotide encoding an outer membrane protein (OMP) of Ehrlichia	
PT	canis or Ehrlichia chaffeensis, useful in preparing a composition for	
PT	diagnosing or preventing E. canis or E. chaffeensis infection.	
XX		
XX		
PS	Disclosure; SEQ ID NO 4; 122pp; English.	
XX		
CC	The invention relates to nucleic acid sequences encoding outer membrane	
CC	proteins (OMP) of Ehrlichia chaffeensis (designated as OMP proteins) and	
CC	Ehrlichia canis (designated as P30F proteins). The OMP polynucleotide is	

```
CC useful in preparing a composition for diagnosing, treating or preventing
CC an infection with Ehrlichia canis or Ehrlichia chaffeensis. The present
CC sequence is the Ehrlichia chaffeensis OMP protein.
XX
SQ Sequence 283 AA;

Query Match      35.4%; Score 507; DB 9; Length 283;
Best Local Similarity 40.3%; Pred. No. 2.7e-45;
Matches 116; Conservative 46; Mismatches 100; Indels 26; Gaps 11;

QY   1 MNVKEPVGVGVTLTSTVFFFISDGAFL-----SDANFSEGRRLGYIGSOYKVGIPNFNSFSA S4
     ||||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB   1 MNVKIIFVSALISLMSILPYQSFDAPVTSNDTGINDSREGFYIVSKYNPSISHFRKFSA 60

QY   55 EET-IPLG----ITKKIFALGLDKSEINTHANSFTSYDPTVA---SSPAGFSGIICYYVNDF 107
     ||||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB   61 EEAPINGNTSIITKVFGLLKKD-GDIAQA SANFNRT-DPALEFQNNLISGFSGSIGAYMDGP 118

QY   108 RVFEFGSYENFERQWYPENSOS-----YKFPALSRNATNSONKFVLNNGVVDKS LN 162
     || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB   119 RIEEAAYQFKDAKN---PDNNDTNSGDYYYKYFGLGREDAIDOKYVVYLKNEGITTFMSLM 175

QY   163 VNVCYDIASGIS I PLAPVMCAGUGADVIK-PLGISLPKFSYQVKFPGVNYPLNNVMTLFEGGG 221
     ||||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB   176 VNTCYDITAEGVFIPIYCACAGUADLIINVPKDNL-KFSYQGKGIGISYPIPTPEVS AFIGG 234

QY   222 YHKVGVGDHRHERVEIAYHTALTALSDVPRTTSSA TLNTDYFGWEIGIFFR 269
     || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB   235 YHGVIIGNPNFKTPV-ITPWVFGAPOTTSAIVTIDTGYEGGEVGVVF 281
```

RESULT 13	
AAV06957	
ID	AAV06957 standard; protein; 285 AA.
XX	AC
XX	AAV06957;
XX	
DT	27-AUG-2003 (revised)
DT	05-JUL-1999 (first entry)
XX	
XX	E. chaffeensis OMP-1Y protein.
XX	
KW	Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;
KW	detection; dog.
XX	
OS	Ehrlichia chaffeensis.
XX	
PN	WO9913720-A1.
XX	
PD	25-MAR-1999.
XX	
PF	18-SEP-1998; 98WO-US019600.
XX	
PK	19-SEP-1997; 97US-0059353P.
XX	
PA	(OHIS ) UNIV OHIO STATE.
XX	
PI	Rikihisa Y, Ohashi N;
XX	
DR	WPI; 1999-254290/21.
DR	N-PSDB; AAIX34757.
XX	
XX	Novel outer membrane proteins from Ehrlichia chaffeensis and Ehrlichia
PT	canis.
XX	
PS	Disclosure: Fig 17B: 55pp; Enlish.

The invention provides isolated outer membrane proteins (OMP) from *Escherichia chaffeensis* and *E. canis*. The *E. chaffeensis* proteins form part of the OMP family and consist of proteins OMP-1, -1(B to Z) shown in CC AY06943-958. The *E. canis* proteins form part of the P30 family and consist of proteins shown in AY06959-970. The proteins and genes are used to detect *E. chaffeensis* in patients and *E. canis* in dogs. (Updated

CC on 27-AUG-2003 to correct OS field.)

XX  
SQ Sequence 285 AA;  
  
Query Match 34.4%; Score 494; DB 2; Length 285;  
Best Local Similarity 38.4%; Pred. No. 6.8e-44;  
Matches 111; Conservative 57; Mismatches 95; Indels 26; Gaps 10;  
  
QY 1 MNYKR--FVGVTLSTFVFLSDGAFSD--ANFSEGRGLYIGSOYKVGIPNFSNFSAE 56  
DB 1 MNNRKSFFIIGASLLASLLTSEASSTGNVSNHTYFKPRLYISGQRPVGVSHFSKFSVKE 60  
  
QY 57 T-----IPGITKKIFALGLDKSEINTHSNFTSRSDPTVYASSPAGSGIIGY-YVNDFRV 109  
DB 61 TNYNTTQLVGLKDDISVIG--NSNITTYTNFNPPIAEFQDNALISFGSAIGLYLSENFR 118  
  
QY 110 EFGSYENPEPERQWYVENS--QSYKFFALSR-----NATNSDN--KFIVLENNGVVDKS 160  
DB 119 EVEASYEEFDVKN---PEGSATDAYRYFALARAMDGTNKSPPDTRKFTVMRNDGLSIS 175  
  
QY 161 LNVNVCYDIASGISPLAPYMCAGVADYIKFLGSLPKFSYQVKGWYVPLNVTMLFGG 220  
DB 176 VMINGCYNFTLDDIPVVPVVCAGIGDFFIEFFNDLHVKAHQKVGISYSISPEVSLFLN 235  
  
QY 221 GYHKVVGDRHERVEIAYHPTALSDVPRTTASATLNTDYFGWEIGFRF 269  
DB 236 GYHKVTGNRFKNLHV-QHVSLSLSDAPKFTSAVATLNVGYFGGEIGVRP 283

RESULT 14

AAU73408  
ID AAU73408 standard; protein; 285 AA.  
XX  
AC AAU73408;  
XX  
DT 12-MAR-2002 (first entry)  
XX  
DE Ehrlichia chaffeensis outer membrane protein P28-9.  
XX  
KW Ehrlichia; outer membrane protein; P28; antibiotic; vaccine.  
XX  
OS Ehrlichia chaffeensis.  
XX  
PN WO200183699-A2.  
XX  
PD 08-NOV-2001.  
XX  
PF 01-MAY-2001; 2001WO-US013997.  
XX  
PR 01-MAY-2000; 2000US-0201035P.  
XX  
PA (RERE-) RES DEV FOUND.  
XX  
PI Walker DH, Yu X;  
XX  
DR WPI; 2002-066527/09.  
XX  
PT Novel Ehrlichia chaffeensis 28-kDa outer membrane protein, designated P28  
PT useful as a vaccine against Ehrlichia chaffeensis.  
XX  
PS Claim 10; Fig 2; 97pp; English.

CC The invention relates to isolated and purified 28-kDa outer membrane  
CC proteins (P28-1 to P28-21) of Ehrlichia chaffeensis. P28 proteins are  
CC encoded by a 28kDa outer membrane protein multigene family. P28 proteins  
CC are useful as a vaccine against E.chaffeensis. DNA encoding P28 is useful  
CC for transfecting a host cell. AAU73400-AAU73420 represent Ehrlichia  
CC chaffeensis P28 outer membrane proteins of the invention  
XX

SQ Sequence 285 AA;

Query Match 34.4%; Score 494; DB 5; Length 285;  
Best Local Similarity 38.4%; Pred. No. 6.8e-44;

Matches 111; Conservative 57; Mismatches 95; Indels 26; Gaps 10;  
  
QY 1 MNYKR--FVGVTLSTFVFLSDGAFSD--ANFSEGRGLYIGSOYKVGIPNFSNFSAE 56  
DB 1 MNNRKSFFIIGASLLASLLTSEASSTGNVSNHTYFKPRLYISGQRPVGVSHFSKFSVKE 60  
  
QY 57 T-----IPGITKKIFALGLDKSEINTHSNFTSRSDPTVYASSPAGSGIIGY-YVNDFRV 109  
DB 61 TNYNTTQLVGLKDDISVIG--NSNITTYTNFNPPIAEFQDNALISFGSAIGLYLSENFR 118  
  
QY 110 EFGSYENPEPERQWYVENS--QSYKFFALSR-----NATNSDN--KFIVLENNGVVDKS 160  
DB 119 EVEASYEEFDVKN---PEGSATDAYRYFALARAMDGTNKSPPDTRKFTVMRNDGLSIS 175  
  
QY 161 LNVNVCYDIASGISPLAPYMCAGVADYIKFLGSLPKFSYQVKGWYVPLNVTMLFGG 220  
DB 176 VMINGCYNFTLDDIPVVPVVCAGIGDFFIEFFNDLHVKAHQKVGISYSISPEVSLFLN 235  
  
QY 221 GYHKVVGDRHERVEIAYHPTALSDVPRTTASATLNTDYFGWEIGFRF 269  
DB 236 GYHKVTGNRFKNLHV-QHVSLSLSDAPKFTSAVATLNVGYFGGEIGVRP 283

RESULT 15

ABG77949  
ID ABG77949 standard; protein; 285 AA.  
XX  
AC ABG77949;  
XX  
DT 15-NOV-2002 (first entry)  
XX  
DE Ehrlichia chaffeensis outer membrane protein (OMP) #15.  
XX  
KW Outer membrane protein; OMP; P30F; ehrlichiosis; infection.  
XX  
OS Ehrlichia chaffeensis.  
XX  
PN US2002120115-A1.  
XX  
PD 29-AUG-2002.  
XX  
PF 28-JAN-2002; 2002US-00059964.  
XX  
PR 19-MAY-1999; 99US-00314701.  
XX  
PA (RIKI/) RIKIHISA Y.  
PA (OHAS/) OHASHI N.  
XX  
PI Rikihisa Y, Ohashi N;  
XX  
DR WPI; 2002-618954/66.  
DR N-PSDB; ABS63290.  
XX  
PT Isolated polynucleotide encoding an outer membrane protein of E.canis or  
PT E.chaffeensis used in the diagnosis of infection.  
XX  
PS Claim 14; Fig 17B; 49pp; English.

CC The invention relates to an isolated polynucleotide encoding an outer  
CC membrane protein (OMP) of Ehrlichia canis or Ehrlichia chaffeensis used  
CC in the diagnosis of infection. An infection such as human ehrlichiosis or  
CC canine ehrlichiosis can be diagnosed by providing a serum sample from the  
CC patient, providing a polypeptide or mixture of polypeptides, contacting  
CC the sample with the polypeptide and assaying for the formation of a  
CC complex between antibodies in the serum sample and the polypeptide, where  
CC formation of a complex is indicative of infection with E. chaffeensis.  
CC This sequence represents an Ehrlichia outer membrane protein of the  
CC invention  
XX

SQ Sequence 285 AA;

Query Match 34.4%; Score 494; DB 5; Length 285;  
Best Local Similarity 38.4%; Pred. No. 6.8e-44;



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 28, 2005, 10:46:11 ; Search time 197 Seconds  
(without alignments)  
7335.826 Million cell updates/sec

Title: US-10-731-554-45  
Perfect score: 813  
Sequencing: 1 atgaattacaaaagattgt.....ttggattagattgcgcta 813

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues  
Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/1 COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5 COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/H COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/PCITUS COMB.seq.\*  
7: /cgn2\_6/ptodata/1/ina/PP COMB.seq.\*  
8: /cgn2\_6/ptodata/1/ina/RE COMB.seq.\*  
9: /cgn2\_6/ptodata/1/ina/backfileseq1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	813	100.0	813	3	US-09-660-587-45
2	813	100.0	813	3	US-09-811-007A-45
3	162.2	20.0	840	3	US-09-314-701-23
4	162.2	20.0	840	3	US-10-314-639-23
5	162.2	20.0	840	3	US-10-059-964A-23
6	150.8	18.5	852	3	US-09-314-701-25
7	150.8	18.5	852	3	US-10-314-639-25
8	150.8	18.5	852	3	US-10-059-964A-25
9	147.2	18.1	840	3	US-09-660-587-41
10	147.2	18.1	840	3	US-09-811-007A-41
11	147.2	18.1	843	3	US-09-314-701-47
12	147.2	18.1	843	3	US-10-314-639-47
13	147.2	18.1	843	3	US-10-059-964A-47
14	146.6	18.0	726	3	US-09-314-701-61
15	146.6	18.0	726	3	US-10-314-639-61
16	146.6	18.0	726	3	US-10-059-964A-61
17	146.4	18.0	849	3	US-09-648-520B-48
18	135	16.6	828	3	US-09-314-701-27
19	135	16.6	828	3	US-10-314-639-27
20	135	16.6	828	3	US-10-059-964A-27
21	133.8	16.5	840	3	US-09-314-701-59
22	133.8	16.5	840	3	US-10-314-639-59
23	133.8	16.5	840	3	US-10-059-964A-59
24	129.6	15.9	843	3	US-09-314-701-5

25	129.6	15.9	843	3	US-10-314-639-5	Sequence 5, Appli
26	129.6	15.9	843	3	US-10-059-964A-5	Sequence 5, Appli
27	126.4	15.5	837	3	US-08-953-326-9	Sequence 9, Appli
28	126.4	15.5	837	3	US-09-314-701-9	Sequence 9, Appli
29	126.4	15.5	837	3	US-09-553-662-9	Sequence 9, Appli
30	126.4	15.5	837	3	US-10-062-994-9	Sequence 9, Appli
31	126.4	15.5	837	3	US-10-314-639-9	Sequence 9, Appli
32	126.4	15.5	837	3	US-10-059-964A-9	Sequence 9, Appli
33	122.4	15.1	867	3	US-09-314-701-31	Sequence 31, Appl
34	122.4	15.1	867	3	US-10-314-639-31	Sequence 31, Appl
35	122.4	15.1	867	3	US-10-059-964A-31	Sequence 10, Appl
36	121.8	15.0	843	3	US-08-953-326-10	Sequence 11, Appl
37	121.8	15.0	843	3	US-09-314-701-11	Sequence 11, Appl
38	121.8	15.0	843	3	US-09-553-662-10	Sequence 10, Appl
39	121.8	15.0	843	3	US-10-062-994-10	Sequence 10, Appl
40	121.8	15.0	843	3	US-10-314-639-11	Sequence 11, Appl
41	121.8	15.0	843	3	US-10-059-964A-11	Sequence 11, Appl
42	120.8	14.9	897	3	US-09-314-701-51	Sequence 51, Appl
43	120.8	14.9	897	3	US-10-314-639-51	Sequence 51, Appl
44	120.8	14.9	897	3	US-10-059-964A-51	Sequence 51, Appl
45	118.2	14.5	852	3	US-09-314-701-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-660-587-45  
; Sequence 45, Application US/09660587  
; Patent No. 6392023  
; GENERAL INFORMATION:  
; APPLICANT: Walker, David H.  
; APPLICANT: McBride, Jere W.  
; APPLICANT: Yu, Xue-Jie  
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein  
; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof  
; FILE REFERENCE: D6152CIP2  
; CURRENT APPLICATION NUMBER: US/09/660,587  
; PRIOR FILING DATE: 2000-09-12  
; PRIOR APPLICATION NUMBER: 09/261,358  
; PRIOR FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 46  
; SEQ ID NO 45  
; LENGTH: 813  
; TYPE: DNA  
; ORGANISM: Ehrlichia canis  
; FEATURE:  
; OTHER INFORMATION: nucleic acid sequence of E. canis p28-9  
US-09-660-587-45

Query Match	100.0%	Score 813	DB 3	Length 813
Best Local Similarity	100.0%	Pred. No. 1.9e-197		
Matches 813	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1	ATGAATACAAAAGATTGTTGAGGTGTTACCGTGTGATACATTGTTTCTTCTATCT	60	
Db	1	ATGAATACAAAAGATTGTTGAGGTGTTACCGTGTGATACATTGTTTCTTCTATCT	60	
Qy	61	GATGGTGCTTTTCTGATGCAAAATTTTCTGAAGGGAGGAGGACCTTTATATAGGTAGT	120	
Db	61	GATGGTGCTTTTCTGATGCAAAATTTTCTGAAGGGAGGAGGACCTTTATATAGGTAGT	120	
Qy	121	CAGTATAAAGTGTGATATCCCAATTTTGTAGTCTTGATAAGTCTGAGATAAATCTCACAGCAAT	180	
Db	121	CAGTATAAAGTGTGATATCCCAATTTTGTAGTCTTGATAAGTCTGAGATAAATCTCACAGCAAT	180	
Qy	181	ATTACAAAAAGATTGTTGCGTTAGGCTTTGAGTCTTGTAGTCTGAGATAAATCTCACAGCAAT	240	
Db	181	ATTACAAAAAGATTGTTGCGTTAGGCTTTGAGTCTTGTAGTCTGAGATAAATCTCACAGCAAT	240	
Qy	241	TTTACAGCATCATGACCTTACTTATGCAAGGAGTTTTCAGGGTTTGTAGTGTATCATTT	300	
Db	241	TTTACAGCATCATGACCTTACTTATGCAAGGAGTTTTCAGGGTTTGTAGTGTATCATTT	300	

```
QY 301 GGATATTATGTTAATGACCTTTAGGCTAGAAATTTGAAGGTTCTTATGAGAAATTTTGAACCT 360
DB |||||||
QY 301 GGATATTATGTTAATGACCTTTAGGCTAGAAATTTGAAGGTTCTTATGAGAAATTTTGAACCT 360
DB |||||||
QY 361 GAAAGACAATGGTACCTCGAATAGCCAAAGCTACAAAATTTTTTGTCTCGAAAT 420
DB |||||||
QY 361 GAAAGACAATGGTACCTCGAATAGCCAAAGCTACAAAATTTTTTGTCTCGAAAT 420
DB |||||||
QY 421 GCTACAAAATAGTGATTAATAGTTTATAGTACTAGAGAAATAACGGCGTTGTCACAAGTCT 480
DB |||||||
QY 421 GCTACAAAATAGTGATTAATAGTTTATAGTACTAGAGAAATAACGGCGTTGTCACAAGTCT 480
DB |||||||
QY 481 CTTAAATGTAATCTTTGTTTATGATATTTGCTAGTGGTAGTATTCTTTTAGCACCTTATATG 540
DB |||||||
QY 481 CTTAAATGTAATCTTTGTTTATGATATTTGCTAGTGGTAGTATTCTTTTAGCACCTTATATG 540
DB |||||||
QY 541 TGTGCTGGTGTGGTGCAGATTATATAAAAGTTTTTTTAGGTATATCATTTGCTAAAGTTTTCT 600
DB |||||||
QY 541 TGTGCTGGTGTGGTGCAGATTATATAAAAGTTTTTTTAGGTATATCATTTGCTAAAGTTTTCT 600
DB |||||||
QY 601 TATCAAGTTAAAGTTTGGTGTCAACTACCTCTAAATGTTTAACTACTATGTTTGGTGGG 660
DB |||||||
QY 601 TATCAAGTTAAAGTTTGGTGTCAACTACCTCTAAATGTTTAACTACTATGTTTGGTGGG 660
DB |||||||
QY 661 GGTATTACCATAAGGTTGTAGGTGATAGCATGAGAGAGTAGAGAAATAGCTTACCATCCT 720
DB |||||||
QY 661 GGTATTACCATAAGGTTGTAGGTGATAGCATGAGAGAGTAGAGAAATAGCTTACCATCCT 720
DB |||||||
QY 721 ACTGCATTATCTGACGTTCTAGAACTACTTCAAGTTCTGCTACTTTTAAATACTGATTAT 780
DB |||||||
QY 721 ACTGCATTATCTGACGTTCTAGAACTACTTCAAGTTCTGCTACTTTTAAATACTGATTAT 780
DB |||||||
QY 781 TTTGGTTGGAGATTGGATTGGCTA 813
DB |||||||
QY 781 TTTGGTTGGAGATTGGATTGGCTA 813
DB |||||||
```

RESULT 2

```
US-09-811-007A-45
; Sequence 45, Application US/09811007A
; Patent No. 660269
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
; FILE REFERENCE: D6152CIP2
; CURRENT APPLICATION NUMBER: US/09/811,007A
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 09/660,587
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 45
; LENGTH: 813
; TYPE: DNA
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: nucleic acid sequence of E. canis p28-9
US-09-811-007A-45
```

```
Query Match 100.0%; Score 813; DB 3; Length 813;
Best Local Similarity 100.0%; Pred. No. 1.9e-197; Indels 0; Gaps 0;
Matches 813; Conservative 0; Mismatches 0;
```

```
QY 1 ATGAATTACAAAAGATTGTTGAGGTGTACCTCAGTACATTTGTTTTTCTTATCT 60
DB |||||||
QY 1 ATGAATTACAAAAGATTGTTGAGGTGTACCTCAGTACATTTGTTTTTCTTATCT 60
DB |||||||
QY 61 GATGGTCTTTTCTGATGCAAAATTTTCTGAAGGAGGAGGACCTTATATAGTAGT 120
DB |||||||
QY 61 GATGGTCTTTTCTGATGCAAAATTTTCTGAAGGAGGAGGACCTTATATAGTAGT 120
DB |||||||
```

```
QY 121 CAGTATAAAGTTGGTATTCCCAATTTTGTAGTAATTTTTCAGCTGAAGAAACAATTCCTGGT 180
DB |||||||
QY 121 CAGTATAAAGTTGGTATTCCCAATTTTGTAGTAATTTTTCAGCTGAAGAAACAATTCCTGGT 180
DB |||||||
QY 181 ATTACAAAAAGATTTTTCGCTTAGTCTTGATAAGTCTGAGATAAAATCTCACAGCAAT 240
DB |||||||
QY 181 ATTACAAAAAGATTTTTCGCTTAGTCTTGATAAGTCTGAGATAAAATCTCACAGCAAT 240
DB |||||||
QY 241 TTTACACGATCATATGACCTTACTTATGCAAGCAGTTTTTCAGGGTGTAGTGGTATCAT 300
DB |||||||
QY 241 TTTACACGATCATATGACCTTACTTATGCAAGCAGTTTTTCAGGGTGTAGTGGTATCAT 300
DB |||||||
QY 301 GGATATTATGTTTAAATGACCTTTAGGGTAGAAATTTGAAAGTCTCTTATGAGAAATTTTCAACCT 360
DB |||||||
QY 301 GGATATTATGTTTAAATGACCTTTAGGGTAGAAATTTGAAAGTCTCTTATGAGAAATTTTCAACCT 360
DB |||||||
QY 361 GAAAGACAATGGTACCTCGAATAGCCAAAGCTACAAAATTTTTTGTCTCGAAAT 420
DB |||||||
QY 361 GAAAGACAATGGTACCTCGAATAGCCAAAGCTACAAAATTTTTTGTCTCGAAAT 420
DB |||||||
QY 421 GCTACAAAATAGTGATTAATAGTTTATAGTACTAGAGAAATAACGGCGTTGTCACAAGTCT 480
DB |||||||
QY 421 GCTACAAAATAGTGATTAATAGTTTATAGTACTAGAGAAATAACGGCGTTGTCACAAGTCT 480
DB |||||||
QY 481 CTTAAATGTAATCTTTGTTTATGATATTTGCTAGTGGTAGTATTCTTTTAGCACCTTATATG 540
DB |||||||
QY 481 CTTAAATGTAATCTTTGTTTATGATATTTGCTAGTGGTAGTATTCTTTTAGCACCTTATATG 540
DB |||||||
QY 541 TGTGCTGGTGTGGTGCAGATTATATAAAAGTTTTTTTAGGTATATCATTTGCTAAAGTTTTCT 600
DB |||||||
QY 541 TGTGCTGGTGTGGTGCAGATTATATAAAAGTTTTTTTAGGTATATCATTTGCTAAAGTTTTCT 600
DB |||||||
QY 601 TATCAAGTTAAAGTTTGGTGTCAACTACCTCTAAATGTTTAACTACTATGTTTGGTGGG 660
DB |||||||
QY 601 TATCAAGTTAAAGTTTGGTGTCAACTACCTCTAAATGTTTAACTACTATGTTTGGTGGG 660
DB |||||||
QY 661 GGTATTACCATAAGGTTGTAGGTGATAGCATGAGAGAGTAGAGAAATAGCTTACCATCCT 720
DB |||||||
QY 661 GGTATTACCATAAGGTTGTAGGTGATAGCATGAGAGAGTAGAGAAATAGCTTACCATCCT 720
DB |||||||
QY 721 ACTGCATTATCTGACGTTCTAGAACTACTTCAAGTTCTGCTACTTTTAAATACTGATTAT 780
DB |||||||
QY 721 ACTGCATTATCTGACGTTCTAGAACTACTTCAAGTTCTGCTACTTTTAAATACTGATTAT 780
DB |||||||
QY 781 TTTGGTTGGAGATTGGATTGGCTA 813
DB |||||||
QY 781 TTTGGTTGGAGATTGGATTGGCTA 813
DB |||||||
```

RESULT 3

```
US-09-314-701-23
; Sequence 23, Application US/09314701
; Patent No. 6544517
; GENERAL INFORMATION:
; APPLICANT: Rikihisa, Yasuko
; APPLICANT: Ohasi, No. 6544517io
; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
; FILE REFERENCE: 22727/04021
; CURRENT APPLICATION NUMBER: US/09/314,701
; CURRENT FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 840
; TYPE: DNA
; ORGANISM: Ehrlichia chaffeensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(840)
US-09-314-701-23
```

Query Match 20.0%; Score 162.2; DB 3; Length 840;  
Best Local Similarity 53.8%; Pred. No. 9.4e-32;  
Matches 432; Conservative 0; Mismatches 353; Indels 18; Gaps 4;

QY 22 GTAGGTGTACGCTGAGTACATTTTGTGTTTCTTATCTGATGGTCTTTTCTGATGCA 81  
DB 34 GTACTTGCACTCTATTATCATCTTATCTATTGAATCCTTTTCAGCTATAAATCATAAT 93

QY 82 AATTTTCTGAAGGAGGAGGAGGACTTTATATAGGTAGTCAGTATAAAGTTGGTATCCC 141  
DB 94 CATACAGGAAATAACACTAGTGTATATATTACAGGCGAGTATAGACCAGGATATCC 153

QY 142 AATTTTGTAGTAATTTTTCAGCTGAGGAGGAGGACTTTATATAGGTAGTCAGTATAAAGTTGGTATCCC 201  
DB 154 CATTTTGTAGCAATTTTCTCAGTAAAGAACTAAATGTTGATACAAATAGGATAT 213

QY 202 TTAGGTCTTGATAAGTCT---GAGATAAATACTACAGCAATTTTACACGATCATATGAC 258  
DB 214 AAAAAAGTGGCTTCTATCGATCCTTAACATTTTCAAACTTTCAGGTCATATACT 273

QY 259 CCTACTTATGCAAGCAGTTTTCAGGGTTTAGTGATCATATGATATTAATGATGAC 318  
DB 274 GTTACATTTCAAGATAATGCTGCTAGTTTCAGTGGAGCAATTTGATATTTTACCOCGAA 333

QY 319 TTT---AGGGTAGAATTTGAAGTTCTTATGAGAAATTTGAACTGAAAGCAATGGTATC 375  
DB 334 AGTCTAAGACTTGAACCTTGAAGGTTCTTACGAAAAATTTGATGTCAAAGATCTCTAAAGAC 393

QY 376 CCTGAGAAATGCAAGCTACAAATTTTGTGTTTGTCTCGAAATGCTACAAATAGTAT 435  
DB 394 TACTCAGCAAAAGATGCTTTTAGGTTTTCAGTGGAGCAATTTGATGTCAAAGATCTCTAAAGAC 486

QY 436 AAT-----AAGTTTATAGTACTAGAGAAATACGGCGTTGTTGCAAGTCTCTTAAT 486  
DB 454 CCTGATGCTCAAAAATATACAGTTATGAAGATAATGGCTTATCTGTTGCATCAATCATG 513

QY 487 GTAAATGTTGTTATGATATTTGCTAGGTAGTATTTCTTTAGCACCTTATATGTTGCT 546  
DB 514 ATCAATGGTTGTTATGATCTATCTTTTAAATAATTTAGTCGTATCACCTTATATATGTCGA 573

QY 547 GGTCTGTCAGAGTATATAAAGTTTTCAGTAAATTTTGTGTTTGTGTTGTTGTTGTTGTT 606  
DB 574 GGTATGAAATTTCTTAAAGCAACTTTCGAGTAGCTACCTTAATGTTGTTGTTGTTGTT 786

QY 751 CTGATGAAATTTCTTAAAGCAACTTTCGAGTAGCTACCTTAATGTTGTTGTTGTTGTT 810

QY 787 TGGGAGATTGGATTAGATTTCG 809  
DB 811 GGTGAAGCTGGAGTAAAGTTTAC 833

## RESULT 4

US-10-314-639-23  
; Sequence 23, Application US/10314639  
; Patent No. 6893640  
; GENERAL INFORMATION:  
; APPLICANT: Rikihisa, Yasuko  
; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia  
; TITLE OF INVENTION: Chaffeensis  
; FILE OF INVENTION: 22727/04021  
; FILE REFERENCE: 22727/04021  
; CURRENT APPLICATION NUMBER: US/10/314,639  
; CURRENT FILING DATE: 2002-12-09

; PRIOR APPLICATION NUMBER: US/09/314,701  
; PRIOR FILING DATE: 1999-05-19  
; NUMBER OF SEQ ID NOS: 66  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 23  
; LENGTH: 840  
; TYPE: DNA  
; ORGANISM: Ehrlichia chaffeensis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(840)  
US-10-314-639-23

Query Match 20.0%; Score 162.2; DB 3; Length 840;  
Best Local Similarity 53.8%; Pred. No. 9.4e-32;  
Matches 432; Conservative 0; Mismatches 353; Indels 18; Gaps 4;

QY 22 GTAGGTGTACGCTGAGTACATTTTGTGTTTCTTATCTGATGGTCTTTTCTGATGCA 81  
DB 34 GTACTTGCACTCTATTATCATCTTATCTATTGAATCCTTTTCAGCTATAAATCATAAT 93

QY 82 AATTTTCTGAAGGAGGAGGAGGACTTTATATAGGTAGTCAGTATAAAGTTGGTATCCC 141  
DB 94 CATACAGGAAATAACACTAGTGTATATATTACAGGCGAGTATAGACCAGGATATCC 153

QY 142 AATTTTGTAGTAATTTTTCAGCTGAGGAGGAGGACTTTATATAGGTAGTCAGTATAAAGTTGGTATCCC 201  
DB 154 CATTTTGTAGCAATTTTCTCAGTAAAGAACTAAATGTTGATACAAATAGGATAT 213

QY 202 TTAGGTCTTGATAAGTCT---GAGATAAATACTACAGCAATTTTACACGATCATATGAC 258  
DB 214 AAAAAAGTGGCTTCTATCGATCCTTAACATTTTCAAACTTTCAGGTCATATACT 273

QY 259 CCTACTTATGCAAGCAGTTTTCAGGGTTTAGTGATCATATGATATTAATGATGAC 318  
DB 274 GTTACATTTCAAGATAATGCTGCTAGTTTCAGTGGAGCAATTTGATATTTTACCOCGAA 333

QY 319 TTT---AGGGTAGAATTTGAAGTTCTTATGAGAAATTTGAACTGAAAGCAATGGTATC 375  
DB 334 AGTCTAAGACTTGAACCTTGAAGGTTCTTACGAAAAATTTGATGTCAAAGATCTCTAAAGAC 393

QY 376 CCTGAGAAATGCAAGCTACAAATTTTGTGTTTGTCTCGAAATGCTACAAATAGTAT 435  
DB 394 TACTCAGCAAAAGATGCTTTTAGGTTTTCAGTGGAGCAATTTGATGTCAAAGATCTCTAAAGAC 486

QY 436 AAT-----AAGTTTATAGTACTAGAGAAATACGGCGTTGTTGCAAGTCTCTTAAT 486  
DB 454 CCTGATGCTCAAAAATATACAGTTATGAAGATAATGGCTTATCTGTTGCATCAATCATG 513

QY 487 GTAAATGTTGTTATGATATTTGCTAGGTAGTATTTCTTTAGCACCTTATATGTTGCT 546  
DB 514 ATCAATGGTTGTTATGATCTATCTTTTAAATAATTTAGTCGTATCACCTTATATATGTCGA 573

QY 547 GGTCTGTCAGAGTATATAAAGTTTTCAGTAAATTTTGTGTTTGTGTTGTTGTTGTTGTT 606  
DB 574 GGTATGAAATTTCTTAAAGCAACTTTCGAGTAGCTACCTTAATGTTGTTGTTGTTGTT 786

QY 607 GTTAAAGTTTGGTGTCAACTACCTCTAAATGTTTAAATACTATGTTGTTGTTGTTGTTGTT 666  
DB 634 GGAAACTAGTATTAGTTACTTCTTCTTAAAGATAATGTTGTTGTTGTTGTTGTTGTT 693

QY 667 TACATAAGGTTGTAGGTAGGATGAGAGTATAGGATAGGATAGGATAGGATAGGATAGGAT 726  
DB 694 TATCATAGGTTATAGGAAATAAATTTTAAATAATTTTAAATAATTTTAAATAATTTTAA 750

QY 727 TTATCTGACGTTCTTAAAGCAACTTTCGAGTAGCTACCTTAATGTTGTTGTTGTTGTT 786  
DB 751 CTGATGAAATTTCTTAAAGCAACTTTCGAGTAGCTACCTTAATGTTGTTGTTGTTGTT 810

QY 787 TGGGAGATTGGATTAGATTTCG 809  
DB 811 GGTGAAGCTGGAGTAAAGTTTAC 833

```
RESULT 5
US-10-059-964A-23
; Sequence 23, Application US/10059964A
; Patent No. 6923963
; GENERAL INFORMATION:
; APPLICANT: RIKIHISA, YASUKO
; APPLICANT: OHASHI, NORIO
; TITLE OF INVENTION: OUTER MEMBRANE PROTEIN OF EHRlichia CANIS AND EHRlichia
; TITLE OF INVENTION: CHAFFEENSIS
; FILE REFERENCE: 22727-04109
; CURRENT APPLICATION NUMBER: US/10/059,964A
; CURRENT FILING DATE: 2002-01-28
; PRIOR APPLICATION NUMBER: 09/314,701
; PRIOR FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: 60/100,843
; PRIOR FILING DATE: 1998-09-18
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 23
; LENGTH: 840
; TYPE: DNA
; ORGANISM: Ehrlichia chaffeensis
US-10-059-964A-23

Query Match      20.0%; Score 162.2; DB 3; Length 840;
Best Local Similarity 53.8%; Pred. No. 9.4e-32;
Matches 432; Conservative 0; Mismatches 353; Indels 18; Gaps 4;

QY      22 GTAGGTGTTACGCTGAGTACATTTGTTTCTTATCTGATGGTCTTTTCTGATGCA 81
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      34 GTACTTGATCTCTATATATCTTCTTATCTGATGATCTTTTCTGATGATATAT 93

QY      82 AATTTTCTGAAGGAGGAGAGGACTTTATATAGTAGTAGTACAGTATTAAGTTGGTATCC 141
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      94 CATACAGGAATAAACAAGTGTGTATATATATACAGGGCAGTATAGCAGGAGTATCC 153

QY      142 AATTTTGTAAATTTTTCAGCTGAAGAAACAATTCCTGGTATTTACAAAAGAAATTTTGG 201
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      154 CATTTTGAACAATTTCTCAGTAAAGAAACTAATGTTGTATACAACTAGTAGGATAT 213

QY      202 TTAGGTCTTGATAAGTCT---GAGATAAAATCTACAGCAAAATTTTACAGCATCATATGAC 258
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      214 AAAAAAGTGGCTCTCTATCGATCTTACACATTTATCAACTTTCAAGGTCATATACT 273

QY      259 CTTACTTATGCAAGCAGTTTTCAGGGTTTTCAGGTATCATTTGATATATGTTAATGAC 318
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      274 GTTACATTTCAAGATAATGCTGCTAGTTTTCAGTGGAGCAATTTGGATATTTCTTACCCGAA 333

QY      319 TTT---AGGGTAGAATTTGAAAGTCTTATGAGAAATTTTGAACCTGAAGAGCAATGGTAC 375
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      334 AGTCTAAGACTTGAACCTTGAAGTCTTACGAAAAATTTGATGTCAAAGATCCTTAAGAC 393

QY      376 CCTGAGAAATAGCCAAAGCTACAAATTTTGTCTCGAAATGCTACAAATAGTGAT 435
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      394 TACTCAGCAAAAGATGCTTTTAGTTTTCCTAGCAGTAAATAGTCTACTAGTT 453

QY      436 AAT-----AAGTTTATAGTACAGAGATAACGGGGTGTGCAAGTCTCTTAAT 486
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      454 CCTGATGCTCAAAAATATACAGTTATGAAGATAATAGGCTTATCTGTTCATCAATCATG 513

QY      487 GTAATGTTTGTATGATATGCTAGTGGTAGTATTCCTTTAGCACCTTATATGTTGCT 546
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      514 ATCAATGGTGTATGATCTATCTTTTAATAATTTAGTCGTATCACCTTATATATGTCGA 573

QY      547 GGTGTTGGTGCAGATATATAAAGTTTTCAGGTATATCATTTGCTTAAGTTTCTTATCAA 606
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      574 GGTATTTGGTGAAGATTTCAATGAAATTTTGTATCTTGCACATTAACCTTCTTATCAA 633

QY      607 GTTAAAGTTGGTGTCAACTACCCCTTAAATGTTTAAATGTTTAACTATGTTTGGTGGGGTTAT 666
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      634 GGAATACTAGGTATTAGTTATTACTTTCTTCTTAAAGTTAAATGTTATTTGCTGGGTGAC 693
```

```
QY      667 TACCATAAGGTTTGTAGGTGATAGGCATGAGAGAGTAGAAATAGCTTACCATCTTACTGCA 726
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      694 TATCATAGAGTTATAGGAATAATTTAAATTT---AATGTTAAACATGTTGTTACA 750

QY      727 TTATCTGACGTTCCCTAGAACTACTTTCAGCTTCTGCTACTTTTAAATPACTGATTTTGGT 786
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      751 CTTGATGAATTTCCCTAAAGCAACTTCTGCAGTAGCTACACTTAATGTTGCTTATTTTGGT 810

QY      787 TGGGAGATTGGATTTAGATTTTGC 809
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      811 GGTGAAGCTGGAGTAAAGTTTAC 833

RESULT 6
US-09-314-701-25
; Sequence 25, Application US/09314701
; Patent No. 6544517
; GENERAL INFORMATION:
; APPLICANT: RIKIHISA, Yasuko
; APPLICANT: Ohashi, No. 6544517io
; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
; TITLE OF INVENTION: Chaffeensis
; FILE REFERENCE: 22727/04021
; CURRENT APPLICATION NUMBER: US/09/314,701
; CURRENT FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Ehrlichia chaffeensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(852)
US-09-314-701-25
```

```
Query Match      18.5%; Score 150.8; DB 3; Length 852;
Best Local Similarity 54.6%; Pred. No. 7.5e-29;
Matches 402; Conservative 0; Mismatches 307; Indels 27; Gaps 4;

QY      96 GAGGAGAGGACTTTATATAGGTAGTACAGTATAAAGTTGGTATTTCCCAATTTTGTAAATTT 155
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      111 GTGCACCTGGCTATATGTGAGTGGACAATAAACCCTACTGTTCTCACTTTTAGTAAATTT 170

QY      156 TTCAGCTGAAGAAACAATTCCTGGTATTTACAAAAGAAATTTTGGTGTAGGCTTTGATAA 215
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      171 TTCACCTTAAAGAAACCTTATCTGCACACTAAAGAGATTATTAGGACTAGCAAAAGATATTAA 230

QY      216 GTCT-----GAGATAAAATCTACAGCAAAATTTTACAGATCATATGACCTTACTTTATGC 269
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      231 GTCTATTACAGATATACAAACAATAAATAATTCACATCTCTTATACACAAAATTTCA 290

QY      270 AAGCAGTTTTCAGGGTTTGTAGTGGTATCATTTGGATATTTATGTTAATGACTTT---AGGGT 326
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      291 AGATAATGCTGTAGTCTTCAGTGCAGCTGTGGATATATTTCCCAAGACAGTCCAAGGCT 350

QY      327 AGAATTTGAAGGTTCTTATGAGAAATTTGAACCTGAAGACAAATGGTACCCTGAGAAATAG 386
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      351 TGAGTAGAATGGTCTTATGAAGAAATTTGACGTTAAATAATCTCGTAAATTTACGTAGTAGA 410

QY      387 CCAAAAGCTACAAAATTTTGTCTCGAAATGCTACAAATAGTGAT----- 435
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      411 TGAAGCTTCAGGTATATTTGCTTTAGCAAGAGGAATTTGATATCTTCAAAAATATCCTGA 470

QY      436 ----AATAAGTTTATAGTACAGAGATAACGGCGTTGTTGACAGTCTCTTTAATGTA 491
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      471 AACAAATAAGTATGTTGTTATAAAGAACAAATGGCTTATCTGCGCATCCATTATAATCAA 530

QY      492 TGTGTTTATGATATTTGCTAGTGGTAGTATTCCTTTAGCACCTTATATGTTGCTGGTGT 551
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      531 TGGCTGTTATGATTTTCTTTTAAACAATTTTAAAGATATCACCTTACATATGCTAGGGTT 590

QY      552 TGGTGCAGATTATATAAAGTTTTTTTAGGTATATCATTTGCTAGTATTTTCTTATCAAGTTAA 611
```

Db 591 TGGTGGGACATTATAGAAATTTTATAGTCTGTAAAGTTTAAATTTGCTTATCAAGGTAA 650  
Qy 612 GTTTGGTGTCAACTACCTCTAAATGTTAACTATGTTGTTGGTGGGGTTATTACCA 671  
Db 651 GGTAGGTATCAGTTATCCATTTCTCTAATATGATTTATATTTGCTGACGGATATTACCA 710  
Qy 672 TAAGTGTAGGTAGGATGAGAGATAGAAATAGCTTACCATCTACTGCATATTC 731  
Db 711 TAAGGTATAGGAAATAAATTTAAACAATTTAAATGTTCAACAGCTTGTAGTCTTAACAG 770  
Qy 732 TGACGTTCTAGAACTACTTTCAGCTTCTGCTACTTTTAAATAGCTATTTTGGTGGGA 791  
Db 771 TCA---TCCTAAGTCTACTTTTGACAGTAGCTACTCTTAATGTTGAGTATTCGGTAGTA 827  
Qy 792 GATTGGATTTAGATTT 807  
Db 828 ATTGGGTAAATTT 843

RESULT 7  
US-10-314-639-25  
; Sequence 25, Application US/10314639  
; Patent No. 6893640  
; GENERAL INFORMATION:  
; APPLICANT: Rikihisa, Yasuko  
; APPLICANT: Ohasi, No. 6893640io  
; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia  
; TITLE OF INVENTION: Chaffeensis  
; FILE REFERENCE: 22727/04021  
; CURRENT APPLICATION NUMBER: US/10/314,639  
; CURRENT FILING DATE: 2002-12-09  
; PRIOR APPLICATION NUMBER: US/09/314,701  
; PRIOR FILING DATE: 1999-05-19  
; NUMBER OF SEQ ID NOS: 66  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 25  
; LENGTH: 852  
; TYPE: DNA  
; ORGANISM: Ehrlichia chaffeensis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(852)  
US-10-314-639-25

Query Match 18.5%; Score 150.8; DB 3; Length 852;  
Best Local Similarity 54.6%; Pred. No. 7.5e-29;  
Matches 402; Conservative 0; Mismatches 307; Indels 27; Gaps 4;

Qy 96 GAGGAGGAGCTTTATATAGGTAGTGCAGTATATAAGTTGGTATCCCAATTTTAGTAATTT 155  
Db 111 GTGCACGTGGCTATATGTCAGTGACAAATATAAACCTACTGTTCTCCTACTTTAGTAATTT 170  
Qy 156 TTCAGCTGAAGAAACAAATTCCTGGTATTAACAAAAGATTTTTCGGTTAGTCTTGATAA 215  
Db 171 TTCACCTTAAGAAACTTATATCTGCACATAAAGAGTTATTAGGACTAGCAAAAAGATATAA 230  
Qy 216 GTCT-----GAGATAAATACTCACAGCAATTTTACACGATCATATGACCCCTACTTATGC 269  
Db 231 GTCTATTACAGATATAACAAATAAATAATTCACATTCCTTATTAACACAAATTTCA 290  
Qy 270 AAGCAGTTTTCAGGGTTTAGTGGTATCATTTGGATATTATGTTAATGACTTTT---AGGGT 326  
Db 291 AGATAATGCTGTTAGCTTTCAGTGCAGCTGTTGGATATATTTCCCAAGACAGTCCAAGGGT 350  
Qy 327 AGAATTTGAAGTCTTATAGAAATTTTGAACCTGAAGCAATAGGTTACCCCTGAGATAG 386  
Db 351 TGAGGTAGAATGGTCTTATGAAGATTTTGACGTTAAAAAATCCTGGTAAATTACGTAGTAAG 410  
Qy 387 CCAAGCTACAAAATTTTGTCTCGAAATGCTACAAATAGTAGAT-----AGGGT 435  
Db 411 TGAAGCCCTCAGGTATATTTGCTTTTAGCAAGGAATTTGATATCTTCAAAAATATCTCGA 470

Qy 436 ----AATAAGTTTATAGTACTAGAGAATAACGGCGTTGTTGACAACTCTCTTAATGTAAA 491  
Db 471 AACAAATAAGTATGTTGTTTATAAGAACAAATGGCTTATCTGTCGCATCCATTAATAATCAA 530  
Qy 492 TGTTTGTATGATATTGCTAGGTAGTATTCCTTTAGCACCTTATATGTTGCTCTGGTGT 551  
Db 531 TGGCTGTTATGATTTTCTTTAAACAATTTAAAGATATCACCTTTACATATGCGTAGGGTT 590  
Qy 552 TGGTGCAGATTATATAAGTTTATAGGTATATCATTTGCCCTAAGTTTCTTATCAAGTTAA 611  
Db 591 TGGTGGGACATTTTATAGAAATTTTATAGTCTGTAAATTTTAAATTTGCTTATCAAGGTAA 650  
Qy 612 GTTTGGTGTCAACTACCTCTAAATGTTAAATATCTATGTTGTTGGTGGGGTTATTACCA 671  
Db 651 GGTAGGTATCAGTTATCCATTTCTCTAATATGATTTATTTGCTGACGGATATTACCA 710  
Qy 672 TAAGTGTAGGTAGGATGAGAGATAGAAATAGCTTACCATCTACTGCATATTC 731  
Db 711 TAAGGTATAGGAAATAAATTTAAACAATTTAAATGTTCAACAGCTTGTAGTCTTAACAG 770  
Qy 732 TGACGTTCTAGAACTACTTTCAGCTTCTGCTACTTTTAAATAGCTATTTTGGTGGGA 791  
Db 771 TCA---TCCTAAGTCTACTTTTGACAGTAGCTACTCTTAATGTTGAGTATTTTCGGTAGTA 827  
Qy 792 GATTGGATTTAGATTT 807  
Db 828 ATTGGGTAAATTT 843

RESULT 8  
US-10-059-964A-25  
; Sequence 25, Application US/10059964A  
; Patent No. 6923963  
; GENERAL INFORMATION:  
; APPLICANT: RIKIHISA, YASUKO  
; APPLICANT: OHASHI, NORIO  
; TITLE OF INVENTION: OUTER MEMBRANE PROTEIN OF EHRlichia CANIS AND EHRlichia  
; TITLE OF INVENTION: CHAFFEENSIS  
; FILE REFERENCE: 22727-04109  
; CURRENT APPLICATION NUMBER: US/10/059,964A  
; CURRENT FILING DATE: 2002-01-28  
; PRIOR APPLICATION NUMBER: 09/314,701  
; PRIOR FILING DATE: 1999-05-19  
; PRIOR APPLICATION NUMBER: 60/100,843  
; PRIOR FILING DATE: 1998-09-18  
; NUMBER OF SEQ ID NOS: 69  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 25  
; LENGTH: 852  
; TYPE: DNA  
; ORGANISM: Ehrlichia chaffeensis  
US-10-059-964A-25

Query Match 18.5%; Score 150.8; DB 3; Length 852;  
Best Local Similarity 54.6%; Pred. No. 7.5e-29;  
Matches 402; Conservative 0; Mismatches 307; Indels 27; Gaps 4;

Qy 96 GAGGAGGAGCTTTATATAGGTAGTGCAGTATATAAGTTGGTATCCCAATTTTAGTAATTT 155  
Db 111 GTGCACGTGGCTATATGTCAGTGACAAATATAAACCTACTGTTCTCCTACTTTAGTAATTT 170  
Qy 156 TTCAGCTGAAGAAACAAATTCCTGGTATTAACAAAAGATTTTTCGGTTAGTCTTGATAA 215  
Db 171 TTCACCTTAAGAAACTTATATCTGCACATAAAGAGTTATTAGGACTAGCAAAAAGATATAA 230  
Qy 216 GTCT-----GAGATAAATACTCACAGCAATTTTACACGATCATATGACCCCTACTTATGC 269  
Db 231 GTCTATTACAGATATAACAAATAAATAATTCACATTCCTTATTAACACAAATTTCA 290  
Qy 270 AAGCAGTTTTCAGGGTTTAGTGGTATCATTTGGATATTATGTTAATGACTTTT---AGGGT 326  
Db 291 AGATAATGCTGTTAGCTTTCAGTGCAGCTGTTGGATATATTTCCCAAGACAGTCCAAGGGT 350





US-10-314-639-47  
; Sequence 47, Application US/10314639  
; Patent No. 6893640  
; GENERAL INFORMATION:  
; APPLICANT: Rikihisa, Yasuko  
; APPLICANT: Ohashi, No. 6893640  
; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia  
; FILE REFERENCE: 22727/04021  
; CURRENT APPLICATION NUMBER: US/10/314,639  
; CURRENT FILING DATE: 2002-12-09  
; PRIOR APPLICATION NUMBER: US/09/314,701  
; PRIOR FILING DATE: 1999-05-19  
; NUMBER OF SEQ ID NOS: 66  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 47  
; LENGTH: 843  
; TYPE: DNA  
; ORGANISM: Ehrlichia canis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(843)  
US-10-314-639-47

Query Match 18.1%; Score 147.2; DB 3; Length 843;  
Best Local Similarity 53.2%; Pred. No. 6.2e-28;  
Matches 445; Conservative 0; Mismatches 358; Indels 33; Gaps 5;

Qy 1 ATGAATTACAAAAGATTGTTGTTAGGTGTTACGCTGAGTACATTTGTTTTTCTTATCT 60  
Db 1 ATGAATTATAAGAAAATTTCTAGTAAGAAGCGGCTTAATCTCAATTAATGTCAATCTTACCA 60  
Qy 61 GATGGTCTTTTTCTGA-----TGCAAAATTTTCTGAAGGAGGAGGACTTTAT 111  
Db 61 TATCAGTCTTTTGCAGATCTGTAGTTCAAGAACTAATGATAACAAAGAGGCTTCTAC 120  
Qy 112 ATAGGTAGTCAGTATAAAGTTGTTATCCCAATTTTGTAAATTTTTCAGCTGAAGAAACA 171  
Db 121 ATTAGTCAAGGTACAACTCCAAAGTATATCACACTTTAGAAAATTTCTCTGCTGAAGAACT 180  
Qy 172 ATTCCTGGTATTACAAAAGATTTTGGGTTAGGTCTTG-----ATAAGTCTGAG 222  
Db 181 CCTATTAAATGGAACAAATTTCTCACTAAAAAGTTTTCGGACTAAAGAAAGATGGTGAT 240  
Qy 223 ATAAATCTACAGCAATTTTACAGCATCATATGACCTACTTAT-----GCAAGCAGT 276  
Db 241 ATACAAAAGAACGATTTTACAGAGTAGTCCAGGCAATGATTTTCAAAATAACTTA 300  
Qy 277 TTTCAGGGTTTATGTTGATCATTTGATTAATGTTAAATGACTTTTAGGGTAGAAATTGAA 336  
Db 301 ATATCAGGATTTTCAGGAAGTATTGTTACTCTATGGACGGACCAAGATAGAACTTGA 360  
Qy 337 GGTTCTTATGAGAAATTTTGAACCTGAAAGCAATGGTACC-----CTGAGAAATAGCCAA 390  
Db 361 GCTGCATATCAACAATTTAATCCAAAAACACCCGATAACAATGATACTGATAATGGTGAA 420  
Qy 391 AGCTACAAATTTTTCGTTTGTCTCGAATAGCTACAAATAGTGAATTAAGTTTATAGTA 450  
Db 421 TACTATAAACATTTTGCATTTATCTCGTAAAGATGCAATGGAAGATCAGCAATATGTAGTA 480  
Qy 451 CTAGAGAAATACCGCGTTTGTGACAAAGTCTCTTAATGTTAAATGTTTGTATGATATGCT 510  
Db 481 CTTAAAATACCGCATTAATTTTATGTCATTGATGTTTAATCTTGTCTATGACATTA 540  
Qy 511 AGTGGTAGTATTCCTTTAGCACTTATATATGTTGCTGGTGGTGGTGCAGATTTATAAAG 570  
Db 541 GCTGAAGAGGATCTTTTCGTACCATATGCAATGATGTCAGGTTATAGGAGCAGATCTTATCACT 600  
Qy 571 TTTTATAGGTATATCATGCTTATAGTTTCTTATCAAGTTTAAGTTGGTGTCAACTACCT 630  
Db 601 ATTTTAAAGACCTCAATCTAAAATTTGCTTTACCAAGGAAAATAGGTATTAGTACCT 660  
Qy 631 CTAATGTTTAATACTATGTTGTTGGGGGTTATTACCAATGAAGTTCTAGTGATAGG 690







GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: December 28, 2005, 10:46:11 ; Search time 4515 Seconds  
(without alignments)  
10235.598 Million cell updates/sec

Title: US-10-731-554-45  
Perfect score: 813  
Sequence: 1 atgaattacaagaattgt.....ttggatttagattggccta 813

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues  
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_in.\*  
3: gb\_env.\*  
4: gb\_on.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pr.\*  
9: gb\_ro.\*  
10: gb\_sts.\*  
11: gb\_sy.\*  
12: gb\_un.\*  
13: gb\_vi.\*  
14: gb\_htg.\*  
15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Description
1	813	100.0	AR437336 Sequence
C 2	813	100.0	AF082744 Ehrlichia
C 3	811.4	99.8	Continuation (13 o
C 4	809.8	99.6	AF078553 Ehrlichia
C 5	574.6	70.7	AF068234 Ehrlichia
C 6	574.6	70.7	U72291 Ehrlichia c
C 7	386.6	47.6	AY652746 Ehrlichia
C 8	386.6	47.6	Continuation (15 o
C 9	380.6	46.8	AV343331 Ehrlichia
C 10	380.6	45.8	Continuation (15 o
C 11	380.6	45.8	Continuation (15 o
C 12	365.6	45.0	AF319940 Ehrlichia
C 13	162.2	20.0	AR303095 Sequence
C 14	162.2	20.0	AR655791 Ehrlichia
C 15	162.2	20.0	U72291 Ehrlichia c
C 16	160.6	19.8	AF393389 Ehrlichia
C 17	160.6	19.8	AF230642 Ehrlichia
C 18	155.6	19.1	AF319940 Ehrlichia

19	150.8	18.5	852	6	AR303096	AR303096 Sequence
20	150.8	18.5	852	6	AR655792	AR655792 Sequence
21	150.8	18.5	21115	1	AY652746	AY652746 Ehrlichia
22	150.8	18.5	99920	1	CR925677_14	Continuation (15 o
23	150.2	18.5	674	1	AB178801	AB178801 Ehrlichia
24	148.8	18.3	674	1	AB178799	AB178799 Ehrlichia
25	148.8	18.3	674	1	AB178800	AB178800 Ehrlichia
26	148.6	18.3	674	1	AB178794	AB178794 Ehrlichia
27	148	18.2	28254	1	AF078553	AF078553 Ehrlichia
28	148	18.2	110000	1	CP000107_12	Continuation (13 o
29	147.8	18.2	674	1	AB178803	AB178803 Ehrlichia
30	147.6	18.2	5828	1	AF479838	AF479838 Ehrlichia
31	147.2	18.1	674	1	AB178802	AB178802 Ehrlichia
32	147.2	18.1	840	6	AR437334	AR437334 Sequence
33	147.2	18.1	843	6	AR303107	AR303107 Sequence
34	147.2	18.1	843	6	AR655803	AR655803 Sequence
35	147.2	18.1	906	1	AF528511	AF528511 Ehrlichia
36	147.2	18.1	906	1	AF528512	AF528512 Ehrlichia
37	147.2	18.1	906	1	AF528514	AF528514 Ehrlichia
38	147.2	18.1	906	1	AF528515	AF528515 Ehrlichia
39	147.2	18.1	6913	1	AF324792	AF324792 Ehrlichia
40	147.2	18.1	11329	1	AF082744	AF082744 Ehrlichia
41	147.2	18.1	110000	1	CP000107_11	Continuation (12 o
42	147	18.1	674	1	AB178798	AB178798 Ehrlichia
43	146.6	18.0	726	6	AR303114	AR303114 Sequence
44	146.6	18.0	726	6	AR655810	AR655810 Sequence
45	146.6	18.0	3507	1	AF125276	AF125276 Cowdria r

ALIGNMENTS

RESULT 1  
AR437336  
LOCUS AR437336 813 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 45 from patent US 6660269.  
ACCESSION AR437336  
VERSION AR437336.1 GI:40202248  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 813)  
AUTHORS Walker,D.H., Yu,X.-J. and McBride,J.W.  
TITLE Homologous 28-Kilodalton immunodominant protein genes of Ehrlichia canis and uses thereof  
JOURNAL Patent: US 6660269-A 45 09-DEC-2003;  
RESEARCH Development Foundation; Carson City, NV  
FEATURES  
Location/Qualifiers  
source  
1..813  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN			
Query Match	Best Local Similarity	Score	Length
Matches	813; Conservative	100.0%;	DB 6; 813; Indels 0; Gaps 0;
Qy	1	ATGAATTACAAAAGATTGTTGTTAGGTGTTTACGCTCAGTACATTTGTTTTTCTTATCT	60
Db	1	ATGAATTACAAAAGATTGTTGTTAGGTGTTTACGCTCAGTACATTTGTTTTTCTTATCT	60
Qy	61	GATGGTGCTTTTCTGATGCAAAATTTTCTGAAGGGAGGAGACCTTTATATAGGTAGT	120
Db	61	GATGGTGCTTTTCTGATGCAAAATTTTCTGAAGGGAGGAGACCTTTATATAGGTAGT	120
Qy	121	CAGTATAAAGTTGGTATTCCTCAATTTTGTAGTAATTTTTCAGCTGAAGAAACAATTCCTGGT	180
Db	121	CAGTATAAAGTTGGTATTCCTCAATTTTGTAGTAATTTTTCAGCTGAAGAAACAATTCCTGGT	180
Qy	181	ATTACAAAAGATTTTTTGGGTTAGGTCTTTGATAAGTCTGAGATAAATACTCACAGCAAT	240
Db	181	ATTACAAAAGATTTTTTGGGTTAGGTCTTTGATAAGTCTGAGATAAATACTCACAGCAAT	240





QY	721	ACTGCATTATCTCAGCGTCTCTAGAACTACTTCAAGCTTCCTGCTACTTTTAATAACTGATTAT	780
Db	10126	ACTGCATTATCTCAGCGTCTCTAGAACTACTTCAAGCTTCCTGCTACTTTTAATAACTGATTAT	10067
QY	781	TTTGGTTGGGAGATTGGATTAGATTGGCGCTA	813
Db	10066	TTTGGTTGGGAGATTGGATTAGATTGGCGCTA	10034
RESULT 3			
CP000107_12/c			
WPCOMMENT			
Sequence split into 14 fragments LOCUS CP000107 Accession CP000107			
Fragment Name Begin End			
CP000107_00 1 110000			
CP000107_01 100001 210000			
CP000107_02 200001 310000			
CP000107_03 300001 410000			
CP000107_04 400001 510000			
CP000107_05 500001 610000			
CP000107_06 600001 710000			
CP000107_07 700001 810000			
CP000107_08 800001 910000			
CP000107_09 900001 1010000			
CP000107_10 1000001 1110000			
CP000107_11 1100001 1210000			
CP000107_12 1200001 1310000			
CP000107_13 1300001 1315030			
Continuation (13 of 14) of CP000107 from base 1200001 (CP000107 Ehrlichia canis str. Jak			
Query Match 99.8%; Score 811.4; DB 1; Length 110000;			
Best Local Similarity 99.9%; Pred. No. 4.6e-138;			
Matches 812; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	1	ATGAATTACAAAAGATTGTGTAGGTGTGTACGCTCAGTACATTTGTTTTTCTTATCT	60
Db	83345	ATGAATTACAAAAGATTGTGTAGGTGTGTACGCTCAGTACATTTGTTTTTCTTATCT	83286
QY	61	GATGGTCTTTTCTGATGCAAAATTTTCTGAAGGAGGAGGAGCTTTATATAGTAGT	120
Db	83285	GATGGTCTTTTCTGATGCAAAATTTTCTGAAGGAGGAGGAGCTTTATATAGTAGT	83226
QY	121	CAGTATAAGCTTGTATTCCTCAATTTAGTAAATTTTTCAGTCAAGAAACAATTCCTGCT	180
Db	83225	CAGTATAAGCTTGTATTCCTCAATTTAGTAAATTTTTCAGTCAAGAAACAATTCCTGCT	83166
QY	181	ATTACAAAAGATTTTTTGGTGTAGGTCTTTGATAAGTCTGAGATAAATACTCACAGCAAT	240
Db	83165	ATTACAAAAGATTTTTTGGTGTAGGTCTTTGATAAGTCTGAGATAAATACTCACAGCAAT	83106
QY	241	TTTACAGCATATAGACCTCTATTGCAAGCAGTTTTCAGCGGTTTAGTGATCATTT	300
Db	83105	TTTACAGCATATAGACCTCTATTGCAAGCAGTTTTCAGCGGTTTAGTGATCATTT	83046
QY	301	GGATATTATCTTAATGACTTTTAGGTAGAAATTTGAAGTTCTTTATGAGAAATTTTGAACCT	360
Db	83045	GGATATTATCTTAATGACTTTTAGGTAGAAATTTGAAGTTCTTTATGAGAAATTTTGAACCT	82986
QY	361	GAAGAACAATGGTACCTCGAATAGCAAGCTACAAAATTTTTTGTCTCGAAAT	420
Db	82985	GAAGAACAATGGTACCTCGAATAGCAAGCTACAAAATTTTTTGTCTCGAAAT	82926
QY	421	GCTACAAATAGTATTAATAGTTTATAGTACTAGAGAAATACCGGTTGTGCAAGTCT	480
Db	82925	GCTACAAATAGTATTAATAGTTTATAGTACTAGAGAAATACCGGTTGTGCAAGTCT	82866
QY	481	CTTAATGTAATGTTTGTATGATATTTAGTGTAGTATTCCTTTAGCACCTTATATG	540
Db	82865	CTTAATGTAATGTTTGTATGATATTTAGTGTAGTATTCCTTTAGCACCTTATATG	82806
QY	541	TGTGCTGGTGTGGTGAGATTATATAAAGTTTTTATAGGTATATCATTTGCCAAGTTTTCT	600
Db			
Db	82805	TGTGCTGGTGTGGTGAGATTATATAAAGTTTTTATAGGTATATCATTTGCCAAGTTTTCT	82746
QY	601	TATCAAGTTAAGTTTTGGTGTCAACTACCTCTTAATATGTTAATACTATCTGTTTGGTGG	660
Db	82745	TATCAAGTTAAGTTTTGGTGTCAACTACCTCTTAATATGTTAATACTATCTGTTTGGTGG	82686
QY	661	GGTATTATACCATTAAGGTTGTAGTGATAGGAGATGAGAGTAGAATAAGCTTACCATCT	720
Db	82685	GGTATTATACCATTAAGGTTGTAGTGATAGGAGATGAGAGTAGAATAAGCTTACCATCT	82626
QY	721	ACTGCATTATCTGACGTTCTCTAGAACTACTTCAAGCTTCCTGCTACTTTTAATAACTGATTAT	780
Db	82625	ACTGCATTATCTGACGTTCTCTAGAACTACTTCAAGCTTCCTGCTACTTTTAATAACTGATTAT	82566
QY	781	TTTGGTTGGGAGATTGGATTAGATTGGCGCTA	813
Db	82565	TTTGGTTGGGAGATTGGATTAGATTGGCGCTA	82533
RESULT 4			
AF078553/c			
LOCUS			
DEFINITION			
1, complete sequence.			
ACCESSION			
AF078553 AF078554 AF078555 AH006958			
VERSION			
AF078553.2 GI:13512584			
KEYWORDS			
SOURCE			
ORGANISM			
Ehrlichia canis			
Ehrlichia canis			
Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;			
Anaplasmataceae; Ehrlichia.			
REFERENCE			
1 (bases 1 to 28254)			
Ohashi,N., Unver,A., Zhi,N. and Rikihisa,Y.			
Cloning and characterization of multigenes encoding the			
immunodominant 30-kilodalton major outer membrane proteins of			
Ehrlichia canis and application of the recombinant protein for			
serodiagnosis			
J. Clin. Microbiol. 36 (9), 2671-2680 (1998)			
JOURNAL			
PUBMED			
9705412			
REFERENCE			
2 (bases 1 to 28254)			
Ohashi,N., Rikihisa,Y. and Unver,A.			
Analysis of transcriptionally active gene clusters of major outer			
membrane protein multigene family in Ehrlichia canis and E.			
chaffeensis			
J. Infect. Immun. 69 (4), 2083-2091 (2001)			
JOURNAL			
PUBMED			
11254561			
REFERENCE			
3 (bases 1 to 28254)			
Ohashi,N., Unver,A., Zhi,N. and Rikihisa,Y.			
Direct Submission			
Submitted (16-JUL-1998) Department of Veterinary Biosciences, The			
Ohio State University, 1925 Coffey Road, Columbus, OH 43210, USA			
JOURNAL			
REFERENCE			
4 (bases 1 to 28254)			
Ohashi,N., Rikihisa,Y. and Unver,A.			
Direct Submission			
Submitted (29-NOV-2000) Department of Veterinary Biosciences, The			
Ohio State University, 1925 Coffey Road, Columbus, OH 43210, USA			
JOURNAL			
AUTHORS			
TITLE			
REMARK			
COMMENT			
On or before Apr 2, 2001 this sequence version replaced gi:3790556,			
gi:3790555, gi:3790558, gi:3790557.			
FEATURES			
source			
1..28254			
/organism="Ehrlichia canis"			
/mol_type="genomic DNA"			
/strain="Oklaoma"			
/db_xref="taxon:944"			
<1..615			
/codon_start=1			
/transl_table=11			
/product="hypothetical transcriptional regulator"			
/protein_id="AAK28679.1"			
/db_xref="GI:13512585"			
/translation="PEYANEIKAHPLIEDLIEKNIQHKFTGEGIRLDVDDYASKNF			
KKEGAESTKELKSATKVRPHFPVDECVGKEIKRQIMRGMSONQLANKLGITTFQOVQKY			

```

EGKTRIVISRLYQLASLVNVEVRDMLKQEDLNKISCDNPITPPHALRDNEKEFLP
EFNDSKIDSKEVLMMVRAYTCIKNEKVRNIYINLVKALSLDNKS"
843..1733
/gene="p30-19"
CDS
843..1733
/gene="p30-19"
/notes="p30 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein P30-19"
/protein_id="AAK28680.1"
/db_xref="GI:13512586"
/translation="MKLLYHLDNIMIKESAIGIVSFPIALFAPNAPSPVPIDFSNES
EMAGFYASQNIIGPFPSPISAKIYTDKSEKELTILFSKEETETIDLKAGDFKRG
QYVYNNRYTGFGAIGYGGGLRVELEGFTFRDVKQKNPDGHRYPALSKDSEI
QNSGSSSSKNDVVMKNGKFNALMFNACDITIGNSSLPVNCACIGIGQGIIRFL
GTVNIHTLFAKGLGLFLISPKTILFANGYVYKAKDNATNLSVQVPEISAAKPHID
PIVYFNADNYGCEVGLRFL"
2491..3318
/gene="p30-18"
CDS
2491..3318
/gene="p30-18"
/notes="p30 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein P30-18"
/protein_id="AAK28681.1"
/db_xref="GI:13512587"
/translation="MNNKLSLLYIALILFTSHVSSALVLDHNLVYGIOYKPARHLL
SNLLIKESQDVVEVLALKYDAIGSPLDSTKEVNNFTIKYNDPHVNNRNLGFSVLGY
YNKFRIESEISHFQIKNEGHRVGEKFKALKAPPSTQGYRHVTLNNGISTY
SALINACYDLVLPAPHNITITVYCLGFDHIDVDFLSKTYTFKFSHQKLGASYPISHRMSV
FTEVYHGLFKGFEQLPLNANTSPPOPPHVHTTASAILSIGYGGSVGKIFIL"
3425..3871
/codon_start=1
/transl_table=11
/product="unknown function U2"
/protein_id="AAK28682.1"
/db_xref="GI:13512588"
/translation="MNNKVLQNEVLLALLPYSLNHSDIDIEYNEQPIKLSRLMKNIT
KSQIKHNYKILLKAKISIGWLFKSDNSQVYVYLRSLSKLVSTIALFDPK
NIFLVYKVLDTSLRNFSNAYYFVKHNTITNCFVSRAYFVFLSCSY"
3882..4832
/gene="p30-17"
CDS
3882..4832
/gene="p30-17"
/notes="p30 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein P30-17"
/protein_id="AAK28683.1"
/db_xref="GI:13512589"
/translation="MNVAKVFLVCILFLPSPFSATNNYFVHEICKSIGHFYGVQ
YKGPDPHNRISADDSFTNLLAISHKDYFLSPSTVEVRLGFLSPQEQNLLHYATGG
STLTLNKDSNKFIPGNFTYDNLGVGGIVGYSINNRLEAEFAIEKFNKAPTQ
NDYETFAIATVYVKGTPKPVHYCKMNTGILSSFLVNTCYDFTLAKKIAPIYCL
GVGDFIDFLQOTLKASQYAGLGAISAPNLTFVVDGSPHGMNMQFGLLVDET
DISVSMPSGDNATVASEBTTMLAKNIMIFLAGSIGIRFIS"
4848..5699
/gene="p30-16"
CDS
4848..5699
/gene="p30-16"
/notes="p30 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein P30-16"
/protein_id="AAK28684.1"
/db_xref="GI:13512590"
/translation="MYKLYLSFIISLAQLLFGSFAFSDIKNNNIHGSYITIKYQPTI
SNFKGHLKTDFTEDPIGDIITAPNTNPDFLKNYFVSLYHKDSYKFFENYDLSGL
DASGLLKVNRILPEGSYKNDTKRLAYHRSRGHKFFAIPRTSNFGVLPNEDNYTV
ANNGISIISNIIINLCSETKFNFTPPICIGIGDFIEIFDMVRKTYQGVGISYP
ITPKLVLSISQYHKVIGNKFKFLPLIQFVALKRTDSDNSPEDKDVLTALLTLDLHFSS
IGLSFIF"
5716..6510
/gene="p30-15"
CDS
5716..6510
/gene="p30-15"
/notes="p30 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein P30-15"
/protein_id="AAK28685.1"
/db_xref="GI:13512591"
/translation="WYNIINYVIKYTIALAFLLPRVSPSILIGNIEKSIKLSVHIN
SOYKPSISOISNVLIOENNSKEKINILNSNTITYNMOLENSTNPFPIIGYFKR
LRPAVEDSEYEHILKNDLSKANLSKYKMYNEDQNFETIATDNKLSITSATVNICY
LIDNNITVPLHLCITAVGICSTGFFNDMRFKLLYQKIGLYLINNSVMLFPVYVHK
VMRNKLLNTQYSVDINAFDAITVLANTDIGYFGSEVGVRFIN"
6525..7394
/gene="p30-14"
CDS
6525..7394
/gene="p30-14"
/notes="p30 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein P30-14"
/protein_id="AAK28686.1"
/db_xref="GI:13512592"
/translation="MIQRLNFIMILAFLLLPFPOSFTLYIHDHEITQNVGLYISQO
YKSPYFFKFLIEENSHKTVLMDLANDVTHVTEYVKMDNTKFTNTPYSAKFRNSLIN
HNSMAYTTFIKNGGISILSWINGCYDIAFSNVKISPYCAGIGGDFITLFTWHIKF
AVQKPGISYLVSPSISIFANGHYHKVMDNVFNKLVHKYIYIQLQDAPITTSARAKLRI
GTFGSEVGVRFPV"
7419..8255
/gene="p30-13"
CDS
7419..8255
/gene="p30-13"
/notes="p30 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein P30-13"
/protein_id="AAK28687.1"
/db_xref="GI:13512593"
/translation="MNNKSLILGTLISLFSLLPIKAPSVINHSDISSNVNGLYFTG
QYRPVSHFGFTVRETNIATQOLVSLNTKNKNHIIERTNPSGLITAKQODNAPAS
GAIGYSYPEGLKPEIEISYEKFGVSKNYSQTNVAFALARTQTSNPSDNKYVNMK
NSGLSVASYMINGCNMSPYLVSPYICAGIGEDFIEFDTLYIKLAYQKLGWNTS
LSSRFNIFADMYHKVIGNQKLNLIHAVALDTPPKVTSIAITLVAYFGGEVGIIRF
IL"
8266..8823
/codon_start=1
/transl_table=11
/product="unknown function U3"
/db_xref="GI:13512594"
/translation="MATYMHKAKIIEPMESPILTDYIILSNVODHNSFTICIT
KLYKLISSITKTLINVKONIAVHNLINLYLNKSKFSQTSQSFMYEYTKINOR
ILNFNSLYPTKONNESYKYTSKSHNFNATYTTTSFIHQHKLKYKISITNOR
NKSTKSPSKILAKNINLNNQENSYE"
8816..9685
/gene="p30-12"
CDS
8816..9685
/gene="p30-12"
/notes="p30 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein P30-12"
/protein_id="AAK28689.1"
/db_xref="GI:13512595"
/translation="MNSKTFSLIGLSILICLAACLPQSPFSSSNVNTYNTWHTGLYIS
GLKPSVSHFGSDFSIKETTYNTNTEALFGLKQDISILRNKETQYNNNFVYPTAKFOD
DPASFYAVGYIANNQPRIEIEGSYEFVKNPNTYIDAHRYIALAREKTSYLLS
PKENKYLKNGGISIVSIINGCYDILSNDKSVICTGFGGDFIEFFSAIRFKFA
YQKIGISYSLSSNIIIFTDGYVHKVINSQFKNLNVHVVNELTTDPKVTSATAFLNI
```

gene		EYFGFGLKFIIP" 9694..10533		ORGANISM		Ehrlichia chaffeensis	
Query Match		99.6%; Score 809.8; DB 1; Length 28254;		REFERENCE		Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;	
Best Local Similarity		99.8%; Pred. No. 1.4e-137;		AUTHORS		Anaplasmataceae; Ehrlichia.	
Matches 811; Conservative 0; Mismatches 2; Indels 0; Gaps 0;				TITLE		1 (bases 1 to 2977)	
QY		1 ATGAATTTACAAAAGATTGTTGTGAGGTGTACGCTCAGTACATTTCTGTTTTTCTTATCT 60		JOURNAL		Yu,X.J., McBride,J.W. and Walker,D.H.	
DB		24807 ATGAATTTACAAAAGATTGTTGTGAGGTGTACGCTCAGTACATTTCTTATCT 24748		REFERENCE		Genetic diversity of the 28-kilodalton outer membrane protein gene in human isolates of Ehrlichia chaffeensis	
QY		61 GATGGTCTTTTCTGCATCAAAATTTTCTGAAGGAGGAGGACTTTATAGTAGT 120		JOURNAL		J. Clin. Microbiol. 37 (4), 1137-1143 (1999)	
DB		24747 GATGGTCTTTTCTGCATCAAAATTTTCTGAAGGAGGAGGACTTTATAGTAGT 24688		REFERENCE		3 (bases 1 to 1268)	
QY		121 CAGTATAAGTTGGTATTTCCCAATTTTAGTAATTTTTCAGCTGAAGAAACAATTCCTGTT 180		AUTHORS		Yu,X.J., McBride,J.W. and Walker,D.H.	
DB		24687 CAGTATAAGTTGGTATTTCCCAATTTTAGTAATTTTTCAGCTGAAGAAACAATTCCTGTT 24628		JOURNAL		Direct Submission	
QY		181 ATTACAAAAAGATTTTTCGTTAGGTCTTGATAAGTCTGAGATAAATACTCACGCAAT 240		REFERENCE		Submitted (26-MAY-1998) Pathology, The University of Texas Medical Branch, 301 Univ. Blvd., Galveston, TX 77555, USA	
DB		24627 ATTACAAAAAGATTTTTCGTTAGGTCTTGATAAGTCTGAGATAAATACTCACGCAAT 24568		AUTHORS		4 (bases 1 to 2977)	
QY		241 TTTACACGATCATPATGACCTACTTATGCAAGCAGTTTTTCGAGGGTTTAGTGGTATCAT 300		JOURNAL		Yu,X.J., McBride,J.W., Zhang,X. and Walker,D.H.	
DB		24567 TTTACACGATCATPATGACCTACTTATGCAAGCAGTTTTTCGAGGGTTTAGTGGTATCAT 24508		REMARK		Direct Submission	
QY		301 GGNATATATGTTAATGACTTTAGGGTAGAATTTGAAGGTTCTTATGAGAAATTTGAACT 360		FEATURES		Submitted (01-FEB-2000) Pathology, University of Texas Medical Branch, 301 Univ. Blvd., Galveston, TX 77555, USA	
DB		24507 GGNATATATGTTAATGACTTTAGGGTAGAATTTGAAGGTTCTTATGAGAAATTTGAACT 24448		COMMENT		Sequence update by submitter	
QY		361 GAAAGCAATGGTACCTGAGATAGCCAAAGCTACAAATTTTTTCTGTTGTCGCAAT 420		LOCATION/Qualifiers		On May 24, 2000 this sequence version replaced gi:3192922.	
DB		24447 GAAAGCAATGGTACCTGAGATAGCCAAAGCTACAAATTTTTTCTGTTGTCGCAAT 24388		1..2977		/organism="Ehrlichia chaffeensis"	
QY		421 GCTCAAAATAGTGATATAAGTTTATAGTACTAGAGAATAACGGCGTTGTCGCAAGTCT 480		/mol_type="genomic DNA"		/db_xref="taxon:945"	
DB		24387 GCTCAAAATAGTGATATAAGTTTATAGTACTAGAGAATAACGGCGTTGTCGCAAGTCT 24328		gene		complement (180..995)	
QY		481 CTTAATGTAATGTTTGTATGATATGCTAGTGGTAGTATTCCTTTAGCACCTTATATG 540		CDS		complement (180..995)	
DB		24327 CTTAATGTAATGTTTGTATGATATGCTAGTGGTAGTATTCCTTTAGCACCTTATATG 24268		/note="component of 28-kDa outer membrane protein		/gene="p28-20"	
QY		541 TGTGCTGGTGTGCTGCAGATTATATAAGTTTATAGTATATCATTTGCTAAGTTTCT 600		multigene family"		/gene="p28-20"	
DB		24267 TGTGCTGGTGTGCTGCAGATTATATAAGTTTATAGTATATCATTTGCTAAGTTTCT 24208		/codon_start=1		complement (183..920)	
QY		601 TATCAAGTTAAAGTTTGGTGTCAACTACCTCTAAATGTTAAATACTATGTTGTTGGTGG 660		/transl_table=11		/gene="p28-20"	
DB		24207 TATCAAGTTAAAGTTTGGTGTCAACTACCTCTAAATGTTAAATACTATGTTGTTGGTGG 24148		/product="p28-20 precursor"		complement (1178..2014)	
QY		661 GGTATTACATAAGGTTGTAGTGATAGGCATGAGAGTAGAATACTGTTACCATCT 720		/protein_id="AAC19134.2"		/gene="p28-21"	
DB		24147 GGTATTACATAAGGTTGTAGTGATAGGCATGAGAGTAGAATACTGTTACCATCT 24088		/db_xref="GI:8052525"		complement (1178..2014)	
QY		721 ACTGCATTATCTCAGCTTCTCAGAACTACTTTCAGCTTCTGCTACTTTAAATACTGATTAT 780		/translation="MNYKKFVVVALATLLSFLPNSFSDANVPGRKGFVVGTOYKV		/gene="p28-21"	
DB		24087 ACTGCATTATCTCAGCTTCTCAGAACTACTTTCAGCTTCTGCTACTTTAAATACTGATTAT 24028		GVNPFNSPASETLPLTKSIFALGLDKSSIDHAGFTQAVNPTVASNFAGPGGVIGY		/note="component of 28-kDa outer membrane protein	
QY		781 TTTGGTTGGGAGATTGGATTTAGATTGGCGTA 813		YVDFRPFEGAYENFERQWYPEGESHKFPALSRSTQDNKFIVLNDGVLDKS		multigene family"	
DB		24027 TTTGGTTGGGAGATTGGATTTAGATTGGCGTA 23995		LNVNFCYDIAGHSIPLADYMCAGGADYIKELGISLPKFSQVRFQVNPVPSVNVNLF		/codon_start=1	
RESULT 5				complement (921..995)		/transl_table=11	
AF068234/c				/gene="p28-20"		/product="p28-21"	
LOCUS				complement (183..920)		/protein_id="AAF71791.1"	
DEFINITION		2977 bp DNA linear BCT 24-MAY-2000		/gene="p28-21"		/db_xref="GI:8052526"	
Ehrlichia chaffeensis 28 kDa outer membrane protein gene cluster,				complement (1178..2014)		/translation="MRYRISVIIIMLLVPCCFSGSLDSDLSNRLKPVFLGISYK	
partial sequence.				/product="p28-21"		LSAPLGSFSGISGETRYINGVKTDRVVGKSDILDLADKAMDKNFNFSSEYVPKYDN	
AF068234 AF230643				/note="component of 28-kDa outer membrane protein		INFGLSFTGFSFRNLRLVEGSKYKFDVIDTRNHLVDNNVYRHIALVRSPPTLYDYF	
AF068234.2 GI:8052524				multigene family"		VLKMDGVFYSTILNICYDFADVTNIVPFCVIGEDIIKIPDSIRFKPSPNSKLGIN	
Ehrlichia chaffeensis				/codon_start=1		YLMNSQDMLLFFDVYVYRVRVGEYNNIPVQYVSLPNPLNISTAAKLDMVEYFGAIGIKV	
SOURCE				/transl_table=11		FV"	



Best Local Similarity 81.7%; Pred. No. 1.4e-94; Matches 664; Conservative 0; Mismatches 149; Indels 0; Gaps 0;			
Qy	1	ATGAATTACAAAAGATTGTTGTAGGTGTACGCTGAGTACATTTGTTTTCTTATCT 60	
Dd	995	ATGAATTACAAAATTTGTTGAGGTGTTGCAATGGCTACATGCTTCTTCTTGCCT 936	
Qy	61	GATGTCGTTTTCTGATGCAAAATTTTCTGAAGGAGGAGAGACCTTTATATAGTAGT 120	
Dd	935	GATAATCTTTTTCTGATGCAAAAGCTTCTGAAGGAGGAAAGGATTTATGTAGGTACT 876	
Qy	121	CAGTATAAAGTTGGTATTTCCCAATTTTGTAGTAAATTTTTCAGCTGAAGAAACAATTCCTGCT 180	
Dd	875	CAATAATAAGTTGGTATTTCCCTAAATTTTCAGTAAATTTTTCAGCTGAAGAAACATTCCTCGGA 816	
Qy	181	ATTACAAAAAGATTTTTCGCTTGTAGGTCTGTGATAAGTCTGAGATAAATCTCACAGCAAT 240	
Dd	815	CTCACAAAAGTATTTTTCGCTTGTAGGCTTGTATAAATCTAGCATTAGTGCACCATGCAGGT 756	
Qy	241	TTTACACGATCATATGACCTTCTATGCAAGCAGTTTTTCAGGGTTTTAGTGGTATCATTT 300	
Dd	755	TTTACACAAGCGTATAATCTTACATATGCAAGCAATTTTGTGGTTTTGTGGTGTATTT 696	
Qy	301	GGATATTATGTTAATGACTTTTAGGTAGAAATTTGAAGTCTTATGAGAAATTTTGAACCT 360	
Dd	695	GGATATTATGTTAATGATTTTAGAGTAGAGTTTGAAGGTGCTTATGAAAATTTTGAACCC 636	
Qy	361	GAAGACAAATGTTACCTCGAGATAGCAAGCTTACAAATTTTGTCTTGTCTCGAAAT 420	
Dd	635	GAAGACAGTGTATCTCTGAAGGAGGAGAAAGCCATTAAGTTTTTGTCTCTATCTCGTGAG 576	
Qy	421	GCTACAAATAGTGATAATAAGTTTATAGTACTAGAGATAACCGCGTTGTGTGCAAGTCT 480	
Dd	575	TCCACTGTTCAAGATAATAAGTTTATAGTGTGTAGAAAATGATGCTGTTATGTATAAATCT 516	
Qy	481	CTTAATGTAATGTTTGTATGATATATGCTAGTGGTAGTATTCCTTTAGCACCTTATATG 540	
Dd	515	CTTAATGTGAACCTTTGTTATGATATGCTCATGCTAGTATTCCTTTAGCACCATATATG 456	
Qy	541	TGTCGTCGTTGTGGCAGATTATATAAAGTTTTTGTAGTATATCATTTGCCCTAAGTTTCT 600	
Dd	455	TGTCGTCGTTGTGGAGCAGATTACATAAATTTTGGGAATATCATTTGCCCTAATTTTCT 396	
Qy	601	TATCAAGTTAAGTTTGGTGTCAACTACCTCTAAATGTTAATACTATGTTGTTGGTGGG 660	
Dd	395	TATCAAGTGAATTTGGGGTTAACTATCTCTGAGTGTAAATGTTATGTTATTTGGTGGGA 336	
Qy	661	GGTTATTACATAAGGTTGTAGGTGATAGGATGAGAGATAGAAATAGCTTACCATCCT 720	
Dd	335	GGTTATTATCATAAGGTTATAGGTAATAGGTATGAAAGATAGAAATAGCTTATCATCCT 276	
Qy	721	ACTGCATTATCTGAGTTCCTAGAACTACTTTCAGCTTCTGCTACTTTAAATACTGATTAT 780	
Dd	275	GCACATTAACGAATGTGCTTAAACTACTTTCAGCATCTGCTACTTTAGATACTGATTAT 216	
Qy	781	TTTGGTTGGGAGATTGGATTTAGATTTTGGCGTA 813	
Dd	215	TTTGGTTGGGAAGTTGGTATGAGATTTTACACTA 183	
RESULT 6			
ECU72291/c			
LOCUS			
DEFINITION	ECU72291	27190 bp DNA linear BCT 02-APR-2001	
ACCESSION	Ehrlichia chaffeensis strain Arkansas major outer membrane protein		
VERSION	Omp-1 multigene cluster, complete sequence.		
KEYWORDS	U72291 AF021338		
SOURCE	U72291.2 GI:13511827		
ORGANISM	Ehrlichia chaffeensis		
Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;			
Anaplasmataceae; Ehrlichia.			
REFERENCE	1 (bases 14844 to 21136; 21479 to 22234)		

Ohashi,N., Zhi,N., Zhang,Y. and Rikihisa,Y. Immunodominant major outer membrane proteins of Ehrlichia chaffeensis are encoded by a polymorphic multigene family Infect. Immun. 66 (1), 132-139 (1998) 9423849	2 (bases 1 to 27190) Ohashi,N., Rikihisa,Y. and Unver,A. Analysis of transcriptionally active gene clusters of major outer membrane protein multigene family in Ehrlichia canis and E. chaffeensis Infect. Immun. 69 (4), 2083-2091 (2001) 11254561	3 (bases 1 to 27190) Ohashi,N., Zhi,N., Zhang,Y. and Rikihisa,Y. Direct Submission Submitted (24-SEP-1996) Department of Veterinary Biosciences, The Ohio State University, 1925 Coffey Rd., Columbus, OH 43210-1093, USA	4 (bases 1 to 27190) Ohashi,N., Zhi,N., Zhang,Y. and Rikihisa,Y. Direct Submission Submitted (27-AUG-1997) Department of Veterinary Biosciences, The Ohio State University, 1925 Coffey Rd., Columbus, OH 43210-1093, USA	5 (bases 1 to 27190) Ohashi,N., Rikihisa,Y. and Unver,A. Direct Submission Submitted (29-NOV-2000) Department of Veterinary Biosciences, The Ohio State University, 1925 Coffey Rd., Columbus, OH 43210-1093, USA	Sequence update by submitter On or before Apr 2, 2001 this sequence version replaced gi:2853584, gi:2853273.	Location/Qualifiers 1. 27190 /organism="Ehrlichia chaffeensis" /mol_type="genomic DNA" /strain="Arkansas" /db_xref="taxon:945" 315..956 /codon_start=1 /transl_table=11 /product="hypothetical transcriptional regulator" /protein_id="AAC02808.2" /db_xref="GI:13511828" /translation="MSTHAKNKINKMEDTDEIKAHDPILAGLIDKNMQKNKNGYVFL DDHKSQPTKEGVSTKDLKNSITKIRPHVDECVGKEIKRQIRMRGMSQNLANKGL ITFOQVQYKEGTNRIVISRLYQLASLVNVEVDIMLKLQEDLKNISCDHTVASTSL KDNSEKFIPEFHDISKIDSKEVLMVMVATYCIKNEKVRNIIYNLVKALSNDKQ" 1136..2032 /gene="omp-1M" 1136..2032 /gene="omp-1M" /note="Omp-1 family member" /codon_start=1 /transl_table=11 /product="major outer membrane protein OMP-1M" /protein_id="AAK28658.1" /db_xref="GI:13511829" /translation="MILINMKLFYHLDNIMTKFSAIGIVLSLVTLFACNVFASPIPI FSNESETAGFYASGHYNIQPFPSPISVKYKIDENTERKELTLFTLKESTDTPTFKQKS BFNDKKGYSPIYRNRYTGFSAIGYSGGLRVELEGATFRFDVKQYKKNQYRYFAL CKKQSIESTDSNGNHHVMMKNGEGRVLSLTFNACYDMIVSNSSILVPSACIGIGGITH FLGGTNIHTLTFKALGLGLFISPKTVIFANGYYVVKTKDNSFTNLISVOYVPLELKEAPH IDPIACFNADNYGGSEVGLRFLI" 2498..3352 /gene="omp-1N" 2498..3352 /gene="omp-1N" /note="Omp-1 family member" /codon_start=1 /transl_table=11 /product="major outer membrane protein OMP-1N"
---	---	--	--	--	--	---

```
/protein_id="AAK28659.1"
/db_xref="GI:13511830"
/translation="MSKRNKRVFLVWMLLETPHISLASVLNDHNSMYGVQYKPAR
QHLSKLLIKESAAANTVEVPLGKDLLDLTGTGNTNFNIKNPYENNRUGFGSIF
GYNNKFRIBELSJETPHIKNGGYKRIDCEKHFAKABISGSNNPANNKYVTLIN
NGISLTSALINCYVDGLKHNIIITYSCILGFGVDITDFLSKYTKTFSSYQGLKASYYV
SQQVSIFRIGYVGLFGKFEKIPVNYPCDYSPTPPNSKPHVHTALAMLSIGYYGG
SIGIKFIP"
CDS
3459..3905
/codon_start=1
/transl_table=11
/product="unknown function UN2"
/protein_id="AAK28660.1"
/db_xref="GI:13511831"
/translation="MNNKSRGQVLLFAAYSIFKYINSCTSKFRLSVKFCFVYNPLKYP
YKSTPLKLGKVFVNTYKSTNYLHRSUFGKFGIIRITNFKISILYNFTNPRVLIIY
YKILNKNKNSINVTLLNNTYTNPTATSFINSVIIGIYPGFVYNY"
3915..4808
/gene="omp-1Q"
/gene="omp-1Q"
3915..4808
/gene="omp-1Q"
/notes="omp-1 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein OMP-1Q"
/protein_id="AAK28661.1"
/db_xref="GI:13511832"
/translation="MSYAKVFLILICLILLVPSLSFAIVNDFLKNIGHFYIGGQYKP
GVPRNRLVTVNNIHELMSEDECKRTIPIHMVQSVAAQGLPPEALEAKLHGGY
LPFTLPYNTYKKNLGGAGVIGYSTHPRVEAPYEFKFNITAGVYLHNKIFYFPA
LATMTDHPHOSAEKDYYMKNTGITLSPFIINACYDPIILKTRNAPYLCIGVGGN
FDFLDQVSKFAYQAKVGISYFVSPNIAFFIDGSHGLHNNQFSDLPVVDYSSGFP
TISAKNANPLTSSIGIRFIS"
4825..5682
/gene="omp-1P"
/gene="omp-1P"
4825..5682
/gene="omp-1P"
/notes="omp-1 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein OMP-1P"
/protein_id="AAK28662.1"
/db_xref="GI:13511833"
/translation="MOKLYISFIILSGLLLPKYVFCMHONNIDGSYVTVIKYQLTTPH
KNFYIKETDPTQEPIDGLAKITANTKEDTLKENFSFSLHCTDSYKSVQNDLLGIGL
SVGLFVKFRIFEGAYKNPNTKRLARYKSGYKIFALPRKSEHGLONTGYTVAK
NNGISIIINILCSSETKYKAFPTYICIGVGGDFIEFDVMRIKRAYQKGVGYPIIT
SKLILSINGQYHKVIGNKPELLPVYQVELKRLVTNKTSKDIDQDVTASLTNLSEHS
SEIGLSFIF"
5695..6513
/gene="omp-1T"
/gene="omp-1T"
5695..6513
/gene="omp-1T"
/notes="omp-1 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein OMP-1T"
/protein_id="AAK28663.1"
/db_xref="GI:13511834"
/translation="MYMYNKKHYCIYIVYVITLFFLLPIESLSALIGNVKDLKVSS
TVSSQYKPSIFHFRNFSIQESHPKKSEEPKIKANLNILKSNAYNLQFQDNTTSF
SYTGIFYSGRLRLAEGCYQEFNVKNSNLSIIISNKHYSRHDENYAITTNKLSIA
SWMTNCTYDISINNTSIVPLCTGIGEDLVGLFNTHFKLAYQGVKGYSLNNKILL
FSDIYHKWGNRPNKLYMQYVADPNISETIPILAKLDIGIFGSEIGIRFMFN"
6529..7416
/gene="omp-1U"
/gene="omp-1U"
6529..7416
/gene="omp-1U"
/notes="omp-1 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein OMP-1U"
/protein_id="AAK28664.1"
```

```
/db_xref="GI:13511835"
/translation="MTKKFNFMVNLITLLELFLPLKSPTTVANNNTITQKVGLYISGO
YKPSIPHPKNSVEENDKVDLIGLTDVTYITHEILPDNTKFNTHYIAKKNNEINF
SSAIGYSGQSPRIEISSYGDFFVYNTKNAVQDVNRYFALVRKNGHYSKSPKPHET
SQPDSNPKSPFYTLMKNGVFAVSIINGCYDFSFNNTTISPVCIGVGGDFIEFFE
VMHIKFAQCSQVIGSYPISPSITIPADAHYHKVINNNFNHLHVKYSYELKNSPTTISA
TAKLNIEYFGCEVGMRFIF"
7438..8277
/gene="omp-1V"
/gene="omp-1V"
7438..8277
/gene="omp-1V"
/notes="omp-1 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein OMP-1V"
/protein_id="AAK28665.1"
/db_xref="GI:13511836"
/translation="MSKKKPTITQVTKKSSIDPNTYSNFOGQPYTVTFODNAASF
QYRPGVSHFNSFKETNVDTIQLVGYKKSASSIDPNTYSNFOGQPYTVTFODNAASF
GAIGVSYSPESIRLEGSYKFDVKDPSAKOAFRFPALARNTSTTVPOAKYVTM
KNGLSVASIMINGCYDLSFNNLVSPYICAGIGEDFIEFEDTLHLKAYQKLGISY
YFFPKINVFAGYTHRVIGNRFPKNLVNHNVTLDFFPKATSAVATLNAVYFGGEAGVK
FTF"
8300..8890
/codon_start=1
/transl_table=11
/product="unknown function UN3"
/protein_id="AAK28666.1"
/db_xref="GI:13511837"
/translation="MIYKEKLTTRYGEVILAYLSFILSTYIFLVLNIIIRYNSLAICVI
SLLRTNIFNVSTKKLIDKDCRDTGSNNCYLKGKLNLOIFGYGFSIRFQNNTLI
IPNDSKCGFYTTLDNPNALHYTYLTGSEYRNFEDILVENIICQCKLLINRNSVLNQ
HNKQTLVLIPIPNAREFSNEIRVNISINKESSYEC"
8880..9731
/gene="omp-1W"
/gene="omp-1W"
8880..9731
/gene="omp-1W"
/notes="omp-1 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein OMP-1W"
/protein_id="AAK28667.1"
/db_xref="GI:13511838"
Query Match 70.7%; Score 574.6; DB 1; Length 27190;
Best Local Similarity 81.7%; Pred. No. 6.5e-95;
Matches 664; Conservative 0; Mismatches 149; Indels 0; Gaps 0;
QY 1 ATGAATTACAAAAGATTTGTTGAGTGTTACGCTGAGTACATTTGTTTTTTTCTTATCT 60
Db 23451 ATGAATTACAAAAGATTTGTTGAGTGTTGATTGGCTACATTTCTTTCTTGCT 23392
QY 61 GATGGTGCTTTTCTGATGCAAAATTTTCTGAAGGAGAGAGAGACTTTATATAGGTAGT 120
Db 23391 GATAATTTCTTTTCTGATGCAAAACGTTCTCTGAAGGAGAGAGAGATTTTATGTAGTACT 23332
QY 121 CAGTATAAGTGGTATTTCCCAATTTTAGTAAATTTTTCAGCTGAAGAACAACTTCTCGT 180
Db 23331 CAATAATAAGTGGTGGTTCCTTAATTTTCAGTAAATTTTTCAGCTGAAGAACAACTTCTCGA 23272
QY 181 ATTACAAAAAGATTTTTCGCGTTAGTCTTGATAAGTCTGAGATAAATACTCACAGCAAT 240
Db 23271 CTCACAAAAGTATTTTTCGTTAGCCCTTGATAAATCTAGCATTAGTACCATGCAGGT 23212
QY 241 TTTACACGATCATGACCTACTTATGCAAGCAGTTTTTGCAGGGTTAGTGGTATCATTT 300
Db 23211 TTTACACAAGCGTATAAATCTACATATGCAAGCAATTTTTCGTTGGTGGTGTATT 23152
QY 301 GGATATTATGTTAATGACTTTTAGGCTAGAAATTTGAAGGTTCTTATGAGAAATTTTGAACCT 360
Db 23151 GGATATTATGTTAATGATTTTAGAGTAGAGTTTGAAGGTGCTTATGAAAAATTTGAACCC 23092
QY 361 GAAAGACAATGGTACCCCTGAGAATAGCCAAAGACTACAAAATTTTTTGTGTTGCTCGAAT 420
```



KRYSFSPWRNDIGISVWMFGCYDLFPFNFKISSYACLGIGDGFIBFFBAMKVKF  
AYQAKGISVTSYSPVNLFPADTVYHKSVGNQFNLRVQVAHTRLRTPFTSAIAKLNI  
GYFGVGVRFFIF"

CDS

7221..8069  
/notes="major antigenic protein 1-8; map1-8"  
/codon\_start=1  
/transl\_table=11

/product="Map1-8"

/protein\_id="AAV73817.1"

/db\_xref="GI:56117752"

/translation="MNKNSIIIGTALTFLAFSPIESFESANOTDETIPITAPISGIY  
FTQYKPGISNFSFNNTYNTQKLRKKDAKESNLGTTTFEDTYSVKFQNNI  
ISPGIGIATSGRIEIEGAYESPDVSPGVSKDNYRYFALARSMTKENKEF  
TWMKNGLSVASIMINGCDEALDFALSPYICAGIGEDFIRFFDALHIKIAVOGKLG  
ISRVSPKISLFDVGYHHIIGNQFNLSVHHAVELSEFPKNSAVATLIDIGLGEV  
GARPIF"

CDS

8080..8931  
/notes="major antigenic protein 1-7; map1-7"  
/codon\_start=1  
/transl\_table=11

/product="Map1-7"

/protein\_id="AAV73818.1"

/db\_xref="GI:56117753"

/translation="MGKFMNYKNTLIGIMLALLPKTSFSESINNDNSONPKLYIS  
GOYKPGVPRFNFSAKETNVTKQKLLGLLSGTSRDKYKDFSNITYIPDFQSSSTGFG  
AAVYISYKGRIRPEIESSEYEEFOVNDHCEGTIBPCRYFALARSVGRGWDDNNYT  
VMKNTGLSTVMFNGCYNIMSKLEISPYICIGIGDPIDFPDATHIKLAVOGKLG  
SVSLDPDVTLFDVGYHRIANNOFNRLNLVQLPVELKYBPKIISATATMNVYTFGEV  
IRFIFNS"

CDS

8951..9850  
/notes="major antigenic protein 1-6; map1-6"  
/codon\_start=1  
/transl\_table=11

/product="Map1-6"

/protein\_id="AAV73819.1"

/db\_xref="GI:56117754"

/translation="MNKRNKLIINTALVFLSFLPHQVLSIPINNSISKYGIYFSG  
SYKLBPFLDNRSIKETTSNTKOVIGLSKKKAENVKDIILGYHAAPNEPYTTFONTA  
SGPSGTAGSYNTKLRPEIEVSEKEDAEIPGSIYDDYEDTYFALARETSHLIG  
ASRRITPKNHYHLMKNGVSIYSFMWNCYQBSTSQSNKILPYICGGVGTDLVQFPN  
KLKLUACQVKLGTSYLSYQOLFANVYHVEIGNFYFNKLKPIRTVLPTNTSTELS  
NVSATSTNISYFSGBIGLRFIL"

CDS

9859..10476  
/notes="major antigenic protein 1-5; map1-5"  
/codon\_start=1  
/transl\_table=11

/product="Map1-5"

/protein\_id="AAV73820.1"

/db\_xref="GI:56117755"

/translation="MIKDSVCLSKVOVILLPLSVLSKAIYNOTELQIFKNSVNLPSVY  
RFYTKLEKNNKELDIKDKYSKDYRHSLSRKINSSSELSINNSKLLONIRSY  
RIPTILDMTRYNYKLPDLGINTLPDGIIRIFQSCYEFKVFYAKIAYNVKVSKKYTIIP  
NIISIFNRYNKKITDNKUNKRNFFNIQKLPKLFHITHTGNTFW"

CDS

10516..11409  
/notes="major antigenic protein 1-4; map1-4"  
/codon\_start=1  
/transl\_table=11

/product="Map1-4"

/protein\_id="AAV73821.1"

/db\_xref="GI:56117756"

/translation="MRIFMKSKIIPTNMVLTFTILSPISISHAEILNKDQKYNIIYLS  
TOYKPAIPVFNFSVEKQNTALLVALTSDINNTNINNEKISLPHQHTTYPHAQFQ  
NSIISPGTIGQVLPKNLRVEIEGYSKSPDKNPGYVDNDVARYVPALARDVKNNSYQ  
PDQNTKNTNLAYYTIMKNYGVISIMSVLLGCCYDIDSVKDKASPPYICLIGVDYDIEF  
FETLHKFPAYQCKVQGISYILPLQVLSLFDADGYHKKVKNQFNKLNITIQVRMLANNPKIT  
YAAATNISYFGAEIGARETF"

CDS

11430..12377  
/notes="major antigenic protein 1-3; map1-3"  
/codon\_start=1  
/transl\_table=11

/product="Map1-3"

/protein\_id="AAV73822.1"

/db\_xref="GI:56117757"

/translation="MTNKLTFGTGLALLCIPNTSPSEIKYNNNTNIQYSIYVSQGY  
KPSVSNFRNFSVKETNTYTKNLIGVKDITSLVHTNNNNKHIVSRNPNPQPTIKATG  
ISNPSFNI PYNPEFQDNI INFSGTIGYQFSKSEIEGYSYKI FDKDPCGGMLYDA  
YRYFALAREMNDTKEPKPYQLDNFPNNFYHTWMKGTGLSISIVMNGCHDPRHVELK  
ISPYICAGVINTIEFFDTSHIKFAYQKIGISYPLSNNIKVFNSNGYHKKVGNKFPKN  
LEVIHVNLANHPWYTSATININIGYFGAEGVIRLGLK"

CDS

12401..13273  
/note="major antigenic protein 1-2; map1-2"  
/codon\_start=1  
/transl\_table=11

/product="Map1-2"

/protein\_id="AAV73823.1"

/db\_xref="GI:56117758"

/translation="MNKNLIRATGIVLISLSPINIPSPSETINNNIKTHYAIYISQYK  
PSVHPFSNFSINETNDIQNLVAAKQNISSLDIDIVLKEPSNPNHPYTFEPDNNISF  
47.6%; Score 386.6; DB 1; Length 21115;  
Best Local Similarity 69.6%; Pred. No. 8.8e-61;  
Matches 578; Conservative 0; Mismatches 229; Indels 24; Gaps 3;

Query Match 47.6%; Score 386.6; DB 1; Length 21115;  
Best Local Similarity 69.6%; Pred. No. 8.8e-61;  
Matches 578; Conservative 0; Mismatches 229; Indels 24; Gaps 3;

QY 1 ATGAATTACAAAGATTTGTTGATGCAAAATTTTCTGAAGGAGGAGGAGGACTTTTATAGGT-- 58  
DB 19213 ATGAATTATAAGAAATTTGTTTAGGTGTACATTAAGTGCAATGTTGTTTCTTACTTA 19154

QY 59 -CTGATGGTCTTTTCTGATGCAAAATTTTCTGAAGGAGGAGGAGGACTTTTATAGGT 117  
DB 19153 CCTCAACGTCTATATCTGATATGATGATGTTTTCAGAAAATAGAAAGTAGGTTTATCGAGT 19094

QY 118 AGTCAGTATAAAGTTGGTATTCCCAATTTTCTAGTAAATTTTTCAGCTGAGAGAAAATTCCT 177  
DB 19093 GTTCAGTATAGGACAGGTATACCAAAATTTTGATAATTTTTCAGCTTCAGAAAATTCCT 19034

QY 178 GGTATTACAAAAGATTTTTCGGTTAGTCTTGA-----TAAGTCTGAGATAATACT 231  
DB 19033 GGTTTAACTAAGGGTGTATTGATTTAGATTAGATTGGATCTAAGTAAATTCGATATACAAA 18974

QY 232 CACAGCAATTTTACACGATCATATGACCTACTTATGCAAGCAGGTTTTGCAGGGTTTACT 291  
DB 18973 CGTGTAAATTTTACAAGATTATATACTACTTATAGTACCAGTTCTACAGGGTAGT 18914

QY 292 GGTATCATTTGGATATATGTTTAATGACTTTAGGGTAGAATTTTGAAGGTTCTTATGAGAA 351  
DB 18913 GGTATGTTGGGTACTATTTTGATAATATAAGAAATGGAATTTTGAACAATCTTTATTTCAAGT 18854

QY 352 TTTGAACCTGAAAGCAATGTTGACCTCGAGATAGCAGAACTACAAATTTTTTGTCTTTG 411  
DB 18853 TTCCGGAATTTGAAGGCGATGTTATCTGAAAGGAGTCAAAGTCAATGTTCTGTCTGTA 18794

QY 412 TCTCGAAATGCTA-----CAAAATAGTGAATAAAGTTTATAGTACTAGAG 456  
DB 18793 TCTCGTAGTGCAACTCCTAGTACTGACTCAAGTAAATAATGATTTTGTAGTATTAGAA 18734

QY 457 AATAACGGGTTGTTGACAAAGTCTCTTAATGTAATAATGTTTGTATGATATTCCTAGTGT 516  
DB 18733 AATAATGGGGTTAAGATTAGAACATTTGAATGTAAAAATTTTGTATGATGTTGCTCACGGA 18674

QY 517 AGTATTCCTTTAGCACCTTATATGCTGCTGTTTGGTGCAGATTATATAAGTTTATTA 576  
DB 18673 AATATTCCTTTAGCACCTTACGTTGCTGTTGTTGGGGGATTACGTAATAATTTATA 18614

QY 577 GGTATATCATTTCCCTCAAGTTTTTCTTATCAAGTTAAAGTTTGGTGTCAACTACCTCTAAAT 636  
DB 18613 GGAGTTTCATTCCTCAAAATTTTCTTATCAGCTTAAATTTGGAGTAATACTACCACCTAAGT 18554

QY 637 GTTAATATCATGTTGTTGGTGGGGGTTATTACCATTAAGTTGTAGGTGATAGGCATGAG 696  
DB 18553 ATTTCGTACTATGCTATTTGGTGGGTGATTATATCAATAAGTATGCGGTAGTAAATATGAT 18494

QY 697 ACAGTAGAAATAGCTTACCATCTACTGCATTATCTGAGCTTCTCTAGAACTACTTCAAGCT 756  
DB 18493 AGGGTAAAGAGTGTTTATCATCTGTTTCAGTTAAATACTGTTTCTTAAATATGACATTTGTT 18434

QY 757 TCTGCTACTTTAAATCACTGATTATTTTGGTGGGAGATTGGATTAGATTTT 807

Db 18433 TCTGCTAAATTAGATATTGATATTTTGGTTGTGAGGTGGGATTAGATTT 18383

RESULT 8  
CR925677 14/c  
WPCOMENT

Sequence split into 15 fragments LOCUS CR925677 Accession CR925677

Fragment Name Begin End

CR925677\_00 1 110000

CR925677\_01 100001 210000

CR925677\_02 200001 310000

CR925677\_03 300001 410000

CR925677\_04 400001 510000

CR925677\_05 500001 610000

CR925677\_06 600001 710000

CR925677\_07 700001 810000

CR925677\_08 800001 910000

CR925677\_09 900001 1010000

CR925677\_10 1000001 1110000

CR925677\_11 1100001 1210000

CR925677\_12 1200001 1310000

CR925677\_13 1300001 1410000

CR925677\_14 1400001 1499920

Continuation (15 of 15) of CR925677 from base 1400001 (CR925677 Ehrlichia ruminantium st

Query Match 47.6%; Score 386.6; DB 1; Length 99920;

Best Local Similarity 69.6%; Pred. No. 5.1e-61;

Matches 578; Conservative 0; Mismatches 229; Indels 24; Gaps 3;

Qy 1 ATGAATTACAAAAGATTGTTGTAGGTGTACCGTGTAGTACATTTGTTTCTTTCTTAT-- 58

Db 57702 ATGAATTATAAGGAATTTGTTTGTAGGTGTACATTAAGTGTGATTTGTTTCTTTCTTACTA 57643

Qy 59 -CTGATGGTCTTTTCTGTAGCAAAATTTTCTGAAGGAGGAGACATTTATATAGGT 117

Db 57642 CCTCAACGTGTATCTGTATGTGATGTGATTTTTCAGAAATAGTAGTGTATGCGAGT 57583

Qy 118 AGTCAGTATAAAGTTGGTATTCCCAATTTTAGTAAATTTTTCAGCTGAAGAAACAATTCCT 177

Db 57582 GTTCAGTATAGGACAGGTATACCAAAATTTTGATAAATTTTCAGCTTCAGAAACAATCCCT 57523

Qy 178 GGTATTAACAAAAGATTTTTCGCTTAGTCTTGA-----TAAGTCTGAGATTAATACT 231

Db 57522 GGTTTAACTAAGGCTTTTATGATTTAGATTTGATCTTAAGTAAATCTGATATTACAAA 57463

Qy 232 CACAGCAATTTTACAGTATATGACCCCTACTTATGACAGCTTTTTCAGGCTTAGT 291

Db 57462 CGTGCTAATTTTACAAGATTATATACTCTACTTATAGTACCAGTTCTACAGGGATAGGT 57403

Qy 292 GGTATCATTTGATATTATGTTAATGACTTTTAGGGTGTAGAAATTTGAAGTTCCTTATGAGAAT 351

Db 57402 GGTATGTTGGGTACTATTTTGATATAATAGATGAATTTGAACATCTTATCAAGT 57343

Qy 352 TTGCACTGAAAGCAATGTGACCTGAGAAATGACCAAGCTCAAAATTTTTCCTTTG 411

Db 57342 TTGCAATTTGAAAGCAGTGTATCTCTGAGGAGTCAAGTCAATAGTTCTGTGCTGTA 57283

Qy 412 TCTCGAATGCTA-----CAATAGTGTATAATAGTTTATAGTACTAGAG 456

Db 57282 TCTCGTAGTCAACTCTCTAGTACTGACTCAAGTATAATATAGTTTGTAGTATTAGAA 57223

Qy 457 AATAACGCGTGTGTGACAAAGTCTCTTAATGTAAATGTTTGTATGATATTGCTAGTGGT 516

Db 57222 AATATGGGTTTAAAGTTAGACATTTGATGTAATTTTGTATGATGTTGCTCACGA 57163

Qy 517 AGTATTCCTTTAGCACTTATATGTCGTGCTGTTGTCGAGATTTATAAAGTTTATA 576

Db 57162 AATATTCCTTTAGCACTTATATGTCGTGCTGTTGTCGAGATTTACGTAATAATTTATA 57103

Qy 577 GGTATATCATTCGCTTAAGTTTCTTATCAAGTAAAGTTTGTGTCACCTACCTTAAT 636

Db 57102 GGAGTTTCATTCGCTTAATTTTCTTATCAGCTTAATTTGAGTAACCTACCCACTAAGT 57043

Qy 637 GTTAATACTATGTTGTTGGTGGGGTTATTACCATAGCTTGTAGTGTATAGGCATGAG 696

Db 57042 ATTCGCTACTATGCTATTTGGTGGGTATTATCATAAAGTGATGGGTAGTAAATGAT 56983

Qy 697 AGAGTAGAATAAGCTTACCATCTACTGCAATTCATGACGTTCTCCTAGAACTACTTCAGCT 756

Db 56982 AGGTTAAAGTGTGTTATCATCTCTGTTCACTTAATGATGTTCTTAAATGACATTTGTT 56923

Qy 757 TCTGCTACTTTAAATACTGATTTATTTTGGTGGGAGATGGATTTAGATTT 807

Db 56922 TCTGCTAATTTAGATATTGATTTATTTGGTGTGAGTTGGGATTAGATTT 56872

RESULT 9  
AY343331/c

LOCUS

DEFINITION

AY343331

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

CDS

misc\_feature

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

CDS

/protein\_id="AA010931.1"  
/db\_xref="GI:38098398"  
/translation="MSYKFLLYMAVILLSPVSLAVNLNENIYKFGVAGIOYKPAKY  
HLSYLDKEDCYNTIDAPAKFSEIKQNIQDINTLATLVSANNPIGYNPHYKNS  
YLGISGALGYHYNGRVESEISSERFLKNEGKILDEHKYFVLARSASGNRIYR  
FSPNENEYILMDGIRSTLSLIPNACYDINIHLITYSCVGFADLVDFLGKYSLK  
PSYQTLGIGSYVSSNIIAIAEGYHGLLSRRFDKIPVNSYAIQSPLSNVDVTASALL  
NIRYGGSIGVRFILGSL"  
3858..4568  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AA010932.1"  
/db\_xref="GI:38098399"  
/translation="MNKKIEKYSLLCGLISLLTSTFNYKPIINVSAYSSISSEFY  
SIIERSNRDIISSIPNPSYQIYSEFKELVKISTNKLIIYH  
PVNNYCLSCNVCINDKNTDNTSTTFQHPQKVLFSNFVITYKSLVDHSLNSNI  
KXFLPSQLYKFFYKRPMTSDPISFNFIHQSKMLVSLFKSVKNNIERTISTVR  
YLNLLTLFRLVNYFKVIL"  
4565..5392  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-12"  
/protein\_id="AA010933.1"  
/db\_xref="GI:38098400"  
/translation="MNYKNIFILFILLPSVSSIASNDTSLSPVYQPYTSVQYKPS  
ISYSPSPSIQENIVEILSKENLSATINNPNKGGQYDIKPISEPNYTKNSP  
LGDTGAIGVKNRYRIELEVFEEFDLKPVEYFFHKDAYKFIILKSTPSHPHHLFKN  
NIDITPVAINCYDIPPKNTKIFPYLCFGAGVDVIDFLDKVHKVSYQAKIGVSYFI  
LNLALFVDSGFSYLSNKFTHIPTINIMDPILPSSSAKFNPNVFLSSSGIRFIH"  
5408..6289  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-11"  
/protein\_id="AA010934.1"  
/db\_xref="GI:38098401"  
/translation="MKKLYLNLFLVLLATCLPKFAFGTFHDTNIAKPIYITVNYOPT  
VSNFRPHIKETNFDTKPIEVNINSTGISSRYFKNREFAYSSHNKHVESYKNLDA  
FTTISIGLSLKNPKEAGSKYKVPDPRNVALQCAHNI PALPRENSGMYPLDRPS  
LRNKVGYTLKNGIATLISNMINLCERQSNFTFYICFGIGDGFIEFDITRIKAA  
YQKIGISYPLTNRGTLISGQYHKVIGNQFKELTQIVELKRLPERQPEYDVTALL  
TLDIYFSGEVLGSL"  
6317..7090  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-10"  
/protein\_id="AA010935.1"  
/db\_xref="GI:38098402"  
/translation="MNNKNSITKVIYITVILSFILLPIQSFALIGNITKNEEYSNVI  
TSOLKPTLVYKDFSLKEINAGHSNDDIITODTKFNHNTSSPSGVCYSIGLRLE  
LEGSHKFTMNSDIIISKI SKYQYSTKAVAATDNYTNNNNITLTSIMVNTCYDIT  
IGNSAVPLYCTGIGGDIININATHRLRAYQKIGISYQALNNFPLFADTYHKIMG  
NKPDIYHNSSNITPMLAKIDIGYFSGVGLRIIPNKL"  
7107..7976  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-9"  
/protein\_id="AA010936.1"  
/db\_xref="GI:38098403"  
/translation="MKKKIINILVILLASLFLSTKSAVSLDNIESPQIYLGTYNK  
LSI SHFTNFSQNTVDIIGLKSVPNTEDDILKAKNFNTQNPPIKNNFTGFSG  
ALGYSGKGLRLEASVODFVKSKNKTNDARYFALVRKDDKAFQPODDVLYR  
KVRSPYSWENDGISIASMENGCVDLPSNFKVSYACIGIGDPIEFEFAMKVP  
AYQAKLGISYISPSVNLFPADTYHKSVGNQFNLRVQYAHTRLRLTPIFTSLAKLNI  
GYFGVGGRFFIP"  
8002..8850  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-8"  
/protein\_id="AA010937.1"  
/db\_xref="GI:38098404"  
/translation="MNNKNSIIIGTALFLLAFSPIESPSANQTDFTIPITAPISGIY  
FTQYKPGISNFSNFSKAKETNTNTQKLVRLKKDAKESNLLGVNTNPEDTYSVKFQNNI

CDS

/protein\_id="AA010931.1"  
/db\_xref="GI:38098398"  
/translation="MSYKFLLYMAVILLSPVSLAVNLNENIYKFGVAGIOYKPAKY  
HLSYLDKEDCYNTIDAPAKFSEIKQNIQDINTLATLVSANNPIGYNPHYKNS  
YLGISGALGYHYNGRVESEISSERFLKNEGKILDEHKYFVLARSASGNRIYR  
FSPNENEYILMDGIRSTLSLIPNACYDINIHLITYSCVGFADLVDFLGKYSLK  
PSYQTLGIGSYVSSNIIAIAEGYHGLLSRRFDKIPVNSYAIQSPLSNVDVTASALL  
NIRYGGSIGVRFILGSL"  
3858..4568  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AA010932.1"  
/db\_xref="GI:38098399"  
/translation="MNKKIEKYSLLCGLISLLTSTFNYKPIINVSAYSSISSEFY  
SIIERSNRDIISSIPNPSYQIYSEFKELVKISTNKLIIYH  
PVNNYCLSCNVCINDKNTDNTSTTFQHPQKVLFSNFVITYKSLVDHSLNSNI  
KXFLPSQLYKFFYKRPMTSDPISFNFIHQSKMLVSLFKSVKNNIERTISTVR  
YLNLLTLFRLVNYFKVIL"  
4565..5392  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-12"  
/protein\_id="AA010933.1"  
/db\_xref="GI:38098400"  
/translation="MNYKNIFILFILLPSVSSIASNDTSLSPVYQPYTSVQYKPS  
ISYSPSPSIQENIVEILSKENLSATINNPNKGGQYDIKPISEPNYTKNSP  
LGDTGAIGVKNRYRIELEVFEEFDLKPVEYFFHKDAYKFIILKSTPSHPHHLFKN  
NIDITPVAINCYDIPPKNTKIFPYLCFGAGVDVIDFLDKVHKVSYQAKIGVSYFI  
LNLALFVDSGFSYLSNKFTHIPTINIMDPILPSSSAKFNPNVFLSSSGIRFIH"  
5408..6289  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-11"  
/protein\_id="AA010934.1"  
/db\_xref="GI:38098401"  
/translation="MKKLYLNLFLVLLATCLPKFAFGTFHDTNIAKPIYITVNYOPT  
VSNFRPHIKETNFDTKPIEVNINSTGISSRYFKNREFAYSSHNKHVESYKNLDA  
FTTISIGLSLKNPKEAGSKYKVPDPRNVALQCAHNI PALPRENSGMYPLDRPS  
LRNKVGYTLKNGIATLISNMINLCERQSNFTFYICFGIGDGFIEFDITRIKAA  
YQKIGISYPLTNRGTLISGQYHKVIGNQFKELTQIVELKRLPERQPEYDVTALL  
TLDIYFSGEVLGSL"  
6317..7090  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-10"  
/protein\_id="AA010935.1"  
/db\_xref="GI:38098402"  
/translation="MNNKNSITKVIYITVILSFILLPIQSFALIGNITKNEEYSNVI  
TSOLKPTLVYKDFSLKEINAGHSNDDIITODTKFNHNTSSPSGVCYSIGLRLE  
LEGSHKFTMNSDIIISKI SKYQYSTKAVAATDNYTNNNNITLTSIMVNTCYDIT  
IGNSAVPLYCTGIGGDIININATHRLRAYQKIGISYQALNNFPLFADTYHKIMG  
NKPDIYHNSSNITPMLAKIDIGYFSGVGLRIIPNKL"  
7107..7976  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-9"  
/protein\_id="AA010936.1"  
/db\_xref="GI:38098403"  
/translation="MKKKIINILVILLASLFLSTKSAVSLDNIESPQIYLGTYNK  
LSI SHFTNFSQNTVDIIGLKSVPNTEDDILKAKNFNTQNPPIKNNFTGFSG  
ALGYSGKGLRLEASVODFVKSKNKTNDARYFALVRKDDKAFQPODDVLYR  
KVRSPYSWENDGISIASMENGCVDLPSNFKVSYACIGIGDPIEFEFAMKVP  
AYQAKLGISYISPSVNLFPADTYHKSVGNQFNLRVQYAHTRLRLTPIFTSLAKLNI  
GYFGVGGRFFIP"  
8002..8850  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-8"  
/protein\_id="AA010937.1"  
/db\_xref="GI:38098404"  
/translation="MNNKNSIIIGTALFLLAFSPIESPSANQTDFTIPITAPISGIY  
FTQYKPGISNFSNFSKAKETNTNTQKLVRLKKDAKESNLLGVNTNPEDTYSVKFQNNI

CDS

/protein\_id="AA010931.1"  
/db\_xref="GI:38098398"  
/translation="MSYKFLLYMAVILLSPVSLAVNLNENIYKFGVAGIOYKPAKY  
HLSYLDKEDCYNTIDAPAKFSEIKQNIQDINTLATLVSANNPIGYNPHYKNS  
YLGISGALGYHYNGRVESEISSERFLKNEGKILDEHKYFVLARSASGNRIYR  
FSPNENEYILMDGIRSTLSLIPNACYDINIHLITYSCVGFADLVDFLGKYSLK  
PSYQTLGIGSYVSSNIIAIAEGYHGLLSRRFDKIPVNSYAIQSPLSNVDVTASALL  
NIRYGGSIGVRFILGSL"  
3858..4568  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AA010932.1"  
/db\_xref="GI:38098399"  
/translation="MNKKIEKYSLLCGLISLLTSTFNYKPIINVSAYSSISSEFY  
SIIERSNRDIISSIPNPSYQIYSEFKELVKISTNKLIIYH  
PVNNYCLSCNVCINDKNTDNTSTTFQHPQKVLFSNFVITYKSLVDHSLNSNI  
KXFLPSQLYKFFYKRPMTSDPISFNFIHQSKMLVSLFKSVKNNIERTISTVR  
YLNLLTLFRLVNYFKVIL"  
4565..5392  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-12"  
/protein\_id="AA010933.1"  
/db\_xref="GI:38098400"  
/translation="MNYKNIFILFILLPSVSSIASNDTSLSPVYQPYTSVQYKPS  
ISYSPSPSIQENIVEILSKENLSATINNPNKGGQYDIKPISEPNYTKNSP  
LGDTGAIGVKNRYRIELEVFEEFDLKPVEYFFHKDAYKFIILKSTPSHPHHLFKN  
NIDITPVAINCYDIPPKNTKIFPYLCFGAGVDVIDFLDKVHKVSYQAKIGVSYFI  
LNLALFVDSGFSYLSNKFTHIPTINIMDPILPSSSAKFNPNVFLSSSGIRFIH"  
5408..6289  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-11"  
/protein\_id="AA010934.1"  
/db\_xref="GI:38098401"  
/translation="MKKLYLNLFLVLLATCLPKFAFGTFHDTNIAKPIYITVNYOPT  
VSNFRPHIKETNFDTKPIEVNINSTGISSRYFKNREFAYSSHNKHVESYKNLDA  
FTTISIGLSLKNPKEAGSKYKVPDPRNVALQCAHNI PALPRENSGMYPLDRPS  
LRNKVGYTLKNGIATLISNMINLCERQSNFTFYICFGIGDGFIEFDITRIKAA  
YQKIGISYPLTNRGTLISGQYHKVIGNQFKELTQIVELKRLPERQPEYDVTALL  
TLDIYFSGEVLGSL"  
6317..7090  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-10"  
/protein\_id="AA010935.1"  
/db\_xref="GI:38098402"  
/translation="MNNKNSITKVIYITVILSFILLPIQSFALIGNITKNEEYSNVI  
TSOLKPTLVYKDFSLKEINAGHSNDDIITODTKFNHNTSSPSGVCYSIGLRLE  
LEGSHKFTMNSDIIISKI SKYQYSTKAVAATDNYTNNNNITLTSIMVNTCYDIT  
IGNSAVPLYCTGIGGDIININATHRLRAYQKIGISYQALNNFPLFADTYHKIMG  
NKPDIYHNSSNITPMLAKIDIGYFSGVGLRIIPNKL"  
7107..7976  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-9"  
/protein\_id="AA010936.1"  
/db\_xref="GI:38098403"  
/translation="MKKKIINILVILLASLFLSTKSAVSLDNIESPQIYLGTYNK  
LSI SHFTNFSQNTVDIIGLKSVPNTEDDILKAKNFNTQNPPIKNNFTGFSG  
ALGYSGKGLRLEASVODFVKSKNKTNDARYFALVRKDDKAFQPODDVLYR  
KVRSPYSWENDGISIASMENGCVDLPSNFKVSYACIGIGDPIEFEFAMKVP  
AYQAKLGISYISPSVNLFPADTYHKSVGNQFNLRVQYAHTRLRLTPIFTSLAKLNI  
GYFGVGGRFFIP"  
8002..8850  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-8"  
/protein\_id="AA010937.1"  
/db\_xref="GI:38098404"  
/translation="MNNKNSIIIGTALFLLAFSPIESPSANQTDFTIPITAPISGIY  
FTQYKPGISNFSNFSKAKETNTNTQKLVRLKKDAKESNLLGVNTNPEDTYSVKFQNNI

CDS

/protein\_id="AA010931.1"  
/db\_xref="GI:38098398"  
/translation="MSYKFLLYMAVILLSPVSLAVNLNENIYKFGVAGIOYKPAKY  
HLSYLDKEDCYNTIDAPAKFSEIKQNIQDINTLATLVSANNPIGYNPHYKNS  
YLGISGALGYHYNGRVESEISSERFLKNEGKILDEHKYFVLARSASGNRIYR  
FSPNENEYILMDGIRSTLSLIPNACYDINIHLITYSCVGFADLVDFLGKYSLK  
PSYQTLGIGSYVSSNIIAIAEGYHGLLSRRFDKIPVNSYAIQSPLSNVDVTASALL  
NIRYGGSIGVRFILGSL"  
3858..4568  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AA010932.1"  
/db\_xref="GI:38098399"  
/translation="MNKKIEKYSLLCGLISLLTSTFNYKPIINVSAYSSISSEFY  
SIIERSNRDIISSIPNPSYQIYSEFKELVKISTNKLIIYH  
PVNNYCLSCNVCINDKNTDNTSTTFQHPQKVLFSNFVITYKSLVDHSLNSNI  
KXFLPSQLYKFFYKRPMTSDPISFNFIHQSKMLVSLFKSVKNNIERTISTVR  
YLNLLTLFRLVNYFKVIL"  
4565..5392  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-12"  
/protein\_id="AA010933.1"  
/db\_xref="GI:38098400"  
/translation="MNYKNIFILFILLPSVSSIASNDTSLSPVYQPYTSVQYKPS  
ISYSPSPSIQENIVEILSKENLSATINNPNKGGQYDIKPISEPNYTKNSP  
LGDTGAIGVKNRYRIELEVFEEFDLKPVEYFFHKDAYKFIILKSTPSHPHHLFKN  
NIDITPVAINCYDIPPKNTKIFPYLCFGAGVDVIDFLDKVHKVSYQAKIGVSYFI  
LNLALFVDSGFSYLSNKFTHIPTINIMDPILPSSSAKFNPNVFLSSSGIRFIH"  
5408..6289  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-11"  
/protein\_id="AA010934.1"  
/db\_xref="GI:38098401"  
/translation="MKKLYLNLFLVLLATCLPKFAFGTFHDTNIAKPIYITVNYOPT  
VSNFRPHIKETNFDTKPIEVNINSTGISSRYFKNREFAYSSHNKHVESYKNLDA  
FTTISIGLSLKNPKEAGSKYKVPDPRNVALQCAHNI PALPRENSGMYPLDRPS  
LRNKVGYTLKNGIATLISNMINLCERQSNFTFYICFGIGDGFIEFDITRIKAA  
YQKIGISYPLTNRGTLISGQYHKVIGNQFKELTQIVELKRLPERQPEYDVTALL  
TLDIYFSGEVLGSL"  
6317..7090  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-10"  
/protein\_id="AA010935.1"  
/db\_xref="GI:38098402"  
/translation="MNNKNSITKVIYITVILSFILLPIQSFALIGNITKNEEYSNVI  
TSOLKPTLVYKDFSLKEINAGHSNDDIITODTKFNHNTSSPSGVCYSIGLRLE  
LEGSHKFTMNSDIIISKI SKYQYSTKAVAATDNYTNNNNITLTSIMVNTCYDIT  
IGNSAVPLYCTGIGGDIININATHRLRAYQKIGISYQALNNFPLFADTYHKIMG  
NKPDIYHNSSNITPMLAKIDIGYFSGVGLRIIPNKL"  
7107..7976  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-9"  
/protein\_id="AA010936.1"  
/db\_xref="GI:38098403"  
/translation="MKKKIINILVILLASLFLSTKSAVSLDNIESPQIYLGTYNK  
LSI SHFTNFSQNTVDIIGLKSVPNTEDDILKAKNFNTQNPPIKNNFTGFSG  
ALGYSGKGLRLEASVODFVKSKNKTNDARYFALVRKDDKAFQPODDVLYR  
KVRSPYSWENDGISIASMENGCVDLPSNFKVSYACIGIGDPIEFEFAMKVP  
AYQAKLGISYISPSVNLFPADTYHKSVGNQFNLRVQYAHTRLRLTPIFTSLAKLNI  
GYFGVGGRFFIP"  
8002..8850  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-8"  
/protein\_id="AA010937.1"  
/db\_xref="GI:38098404"  
/translation="MNNKNSIIIGTALFLLAFSPIESPSANQTDFTIPITAPISGIY  
FTQYKPGISNFSNFSKAKETNTNTQKLVRLKKDAKESNLLGVNTNPEDTYSVKFQNNI

CDS

/protein\_id="AA010931.1"  
/db\_xref="GI:38098398"  
/translation="MSYKFLLYMAVILLSPVSLAVNLNENIYKFGVAGIOYKPAKY  
HLSYLDKEDCYNTIDAPAKFSEIKQNIQDINTLATLVSANNPIGYNPHYKNS  
YLGISGALGYHYNGRVESEISSERFLKNEGKILDEHKYFVLARSASGNRIYR  
FSPNENEYILMDGIRSTLSLIPNACYDINIHLITYSCVGFADLVDFLGKYSLK  
PSYQTLGIGSYVSSNIIAIAEGYHGLLSRRFDKIPVNSYAIQSPLSNVDVTASALL  
NIRYGGSIGVRFILGSL"  
3858..4568  
/codon\_start=1  
/transl\_table=11  
/product="unknown"  
/protein\_id="AA010932.1"  
/db\_xref="GI:38098399"  
/translation="MNKKIEKYSLLCGLISLLTSTFNYKPIINVSAYSSISSEFY  
SIIERSNRDIISSIPNPSYQIYSEFKELVKISTNKLIIYH  
PVNNYCLSCNVCINDKNTDNTSTTFQHPQKVLFSNFVITYKSLVDHSLNSNI  
KXFLPSQLYKFFYKRPMTSDPISFNFIHQSKMLVSLFKSVKNNIERTISTVR  
YLNLLTLFRLVNYFKVIL"  
4565..5392  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-12"  
/protein\_id="AA010933.1"  
/db\_xref="GI:38098400"  
/translation="MNYKNIFILFILLPSVSSIASNDTSLSPVYQPYTSVQYKPS  
ISYSPSPSIQENIVEILSKENLSATINNPNKGGQYDIKPISEPNYTKNSP  
LGDTGAIGVKNRYRIELEVFEEFDLKPVEYFFHKDAYKFIILKSTPSHPHHLFKN  
NIDITPVAINCYDIPPKNTKIFPYLCFGAGVDVIDFLDKVHKVSYQAKIGVSYFI  
LNLALFVDSGFSYLSNKFTHIPTINIMDPILPSSSAKFNPNVFLSSSGIRFIH"  
5408..6289  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-11"  
/protein\_id="AA010934.1"  
/db\_xref="GI:38098401"  
/translation="MKKLYLNLFLVLLATCLPKFAFGTFHDTNIAKPIYITVNYOPT  
VSNFRPHIKETNFDTKPIEVNINSTGISSRYFKNREFAYSSHNKHVESYKNLDA  
FTTISIGLSLKNPKEAGSKYKVPDPRNVALQCAHNI PALPRENSGMYPLDRPS  
LRNKVGYTLKNGIATLISNMINLCERQSNFTFYICFGIGDGFIEFDITRIKAA  
YQKIGISYPLTNRGTLISGQYHKVIGNQFKELTQIVELKRLPERQPEYDVTALL  
TLDIYFSGEVLGSL"  
6317..7090  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-10"  
/protein\_id="AA010935.1"  
/db\_xref="GI:38098402"  
/translation="MNNKNSITKVIYITVILSFILLPIQSFALIGNITKNEEYSNVI  
TSOLKPTLVYKDFSLKEINAGHSNDDIITODTKFNHNTSSPSGVCYSIGLRLE  
LEGSHKFTMNSDIIISKI SKYQYSTKAVAATDNYTNNNNITLTSIMVNTCYDIT  
IGNSAVPLYCTGIGGDIININATHRLRAYQKIGISYQALNNFPLFADTYHKIMG  
NKPDIYHNSSNITPMLAKIDIGYFSGVGLRIIPNKL"  
7107..7976  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-9"  
/protein\_id="AA010936.1"  
/db\_xref="GI:38098403"  
/translation="MKKKIINILVILLASLFLSTKSAVSLDNIESPQIYLGTYNK  
LSI SHFTNFSQNTVDIIGLKSVPNTEDDILKAKNFNTQNPPIKNNFTGFSG  
ALGYSGKGLRLEASVODFVKSKNKTNDARYFALVRKDDKAFQPODDVLYR  
KVRSPYSWENDGISIASMENGCVDLPSNFKVSYACIGIGDPIEFEFAMKVP  
AYQAKLGISYISPSVNLFPADTYHKSVGNQFNLRVQYAHTRLRLTPIFTSLAKLNI  
GYFGVGGRFFIP"  
8002..8850  
/codon\_start=1  
/transl\_table=11  
/product="MAPI-8"  
/protein\_id="AA010937.1"  
/db\_xref="GI:38098404"  
/translation="MNNKNSIIIGTALFLLAFSPIESPSANQTDFTIPITAPISGIY  
FTQYKPGISNFSNFSKAKETNTNTQKLVRLKKDAKESNLLGVNTNPEDTYSVKFQNNI

Matches 578; Conservative 0; Mismatches 229; Indels 30; Gaps 3;			
Qy	1	ATGAATTACAAAGATTGTTGTTAGGTGTTACCGCTGAGTACATTTGTTTTTCTTAT-- 58	
Db	20028	ATGAATTATAAGGAATTTGTTTAAAGTGTACATTAAAGTCAATTTGTTTCTTACTA 19969	
Qy	59	-CTGATCGTCTTTTCTGATGCAAAATTTTCTGAAGGGAGGAGACCTTATATAGGT 117	
Db	19968	CCTGAACGTGCTATCTCGATATGGAATTTTTCAGAAAAATAGAAGTAGTTTATGACAGT 19909	
Qy	118	AGTCAGTATAAAGTTGGTATTTCCCAATTTTAGTAAATTTTTCAGCTGAAGAAACAATTCCT 177	
Db	19908	GTTTCAGTATAGGACAGGTATACCAAAATTTTGATPAATTTTTCAGCTTCAGAAACAATCCCT 19849	
Qy	178	GGTATTAACAAAAAGATTTTTCGCTTAGGTCTTGA-----TAAGTCTGAGATAAATPACT 231	
Db	19848	GGTTTAACTAAGGGTGTATGGAATTTAGATTTGGATCTAAGTAAATCTGATATTAACAAA 19789	
Qy	232	CACAGCAATTTTACACGATCATATGACCTACTTATGCAAGCAGTTTTCAGGGTTAGT 291	
Db	19788	CGTGCTAATTTTACAGATTTATATCTCTACTTATAGTACCAGTCTACAGGGTAGGT 19729	
Qy	292	GGTATCATTTGATATTAATGACTTTTAGGGTAGAAATTTGAAGTCTCTTATGAGAAAT 351	
Db	19728	GGTATGTTGGTACTATTTTGATTAATATAGAAATGGAATTTGAAACATCTTATCAAGT 19669	
Qy	352	TTTGAACCTGAAAGCAATGGTACCTTGAGAATAGCCAAAGCTACAAATTTTTCGTTTG 411	
Db	19668	TTCGGAATTTGAAGGCAGTGTATCTCTGAAGGGAGTCAAAGTCAATAGTTTTCGCTGTA 19609	
Qy	412	TCTCGAAATGCTA-----CAAAATAGTGAATAATAAGTTTATAGTA 450	
Db	19608	TCTCGTCAAGATAATGCGAGCTCTTAATGACTCAAGTAAATAATATGATTTTGTAGTG 19549	
Qy	451	CTAGAGAAATACGGCGTTTGTGCAAGTCTCTTAATGTAATGTTTCTGTTATGATATTGCT 510	
Db	19548	TTGGAAATATATGGGTTTAAGATTAGACATTTGCATGTAATTTTGTATGATGTTGCT 19489	
Qy	511	AGTGGTAGTATTCCTTTAGCACCTTATATGTGCTGCTGTTGGTGAGATATATAAAG 570	
Db	19488	CACGGAATATTCCTTTAGCACCTTACGTGTGCTGCTGATTTGGTGGGATTTATGAANA 19429	
Qy	571	TTTTTAGGTATATCAATTTGCTTAAGTTTCTTATCAAGTTAAGTTTGGTGCACTACCT 630	
Db	19428	TTTATAGGAGTTTCATTGCTTAAATTTTCTTATCAGCTTAAATTTGGAGTAAACTACCCA 19369	
Qy	631	CTAAATGTTAATCTATGTTTGTGGGGGTTTATACCATAGGTTGTAGGTATAGG 690	
Db	19368	CTAAGTATTCGTACTATTTGGTGGGTTATTTATCATAAAGTTATGGGTAGTAA 19309	
Qy	691	CATGAGAGAGTAGAAATAGCTTACCCTACTGCTATTCATTTCTGACGTTCTAGAACTPACT 750	
Db	19308	TATGATAGGGTGAAGTGGTTTATCATCTCTGTTCAAGTTAACTGTTCTTAAATGACA 19249	
Qy	751	TCAGCTTCTGCTACTTTTAAATPACTGATATTTTGGTTGGGAGATGGAATTTAGATT 807	
Db	19248	TTTGTCTGCTAATTTAGATATTTGATTTTGGTTGGTGGGTTATTTTGGTTGGGATTAGATT 19192	
RESULT 10			
CR767821_14/c			
WPCOMMENT			
Sequence split into 16 fragments LOCUS CR767821 Accession CR767821			
Fragment Name	Begin	End	
CR767821_00	1	110000	
CR767821_01	100001	210000	
CR767821_02	200001	310000	
CR767821_03	300001	410000	
CR767821_04	400001	510000	
CR767821_05	500001	610000	
CR767821_06	600001	710000	
CR767821_07	700001	810000	
CR767821_08	800001	910000	

CR767821_09	900001	1010000	
CR767821_10	100001	1110000	
CR767821_11	1100001	1210000	
CR767821_12	1200001	1310000	
CR767821_13	1300001	1410000	
CR767821_14	1400001	1510000	
CR767821_15	1500001	1516355	
Continuation (15 of 16) of CR767821 from base 1400001 (CR767821 Ehrlichia ruminantium str			
Query Match 46.8%; Score 380.6; DB 1; Length 110000;			
Best Local Similarity 69.1%; Pred. No. 6.1e-60;			
Matches 578; Conservative 0; Mismatches 229; Indels 30; Gaps 3;			
Qy	1	ATGAATTACAAAGATTGTTGTTAGGTGTTACGCTGAGTACATTTGTTTTTCTTAT-- 58	
Db	93551	ATGAATTATAAGGAATTTGTTTTAGGTGTTACATTAAAGTGCATTTGTTTCTTACTTA 93492	
Qy	59	-CTGATGGTGCCTTTTCTGATGCAAAATTTTCTGAAGGGAGGAGAGGACTTTTATATAGGT 117	
Db	93491	CCTGAACGTGCTATATCTGATATGGAATTTTTCAGAAAAATAGAAGTAGGTTTTATGACAGT 93432	
Qy	118	AGTCAGTATAAAGTTGGTATTTCCCAATTTTAGTAAATTTTTCAGCTGAAGAAACAATTCCT 177	
Db	93431	GTTTCAGTATAGGACAGGTATACCAAAATTTTGATAATTTTTCAGCTTCAGAAACAATCCCT 93372	
Qy	178	GGTATTAACAAAAAGATTTTTCGCTTAGGTCTTGA-----TAAGTCTGAGATAAATACT 231	
Db	93371	GGTTTAACTAAGGGTGTTTATGGAATTTGGATTTTGGATCTAAGTAAATCTGATATTAACAAA 93312	
Qy	232	CACAGCAATTTTACACGATCATATGACCTACTTATGCAAGCAGTTTTCAGGGTTTAGT 291	
Db	93311	CGTGCTAATTTTACAGATTTATATCTCTTATAGTACCAGTCTACAGGGATAGGT 93252	
Qy	292	GGTATCATTTGGATTTATGTTAATGACTTTTAGGGTAGAATTTTGAAGTCTCTTATGAGAAT 351	
Db	93251	GGTATGTTGGTACTATTTTGTATATAAGAAATGGAATTTTGAACAATCTTATTTCAAGT 93192	
Qy	352	TTTGAACCTGAAAGCAATGGTACCTTGAGAAATAGCCAAAGCTACAAATTTTTCGTTTG 411	
Db	93191	TTCGGAATTTGAAGGCAGTGGTATCTCTGAAGGGAGTCAAAGTCAATAAGTTTTCGCTGTA 93132	
Qy	412	TCTCGAAATGCTA-----CAAAATAGTGAATAAAGTTTATAGTA 450	
Db	93131	TCTCGTCAAGATAATGCGAGCTCTTAATGACTCAAGTAAATAATATGATTTGTAGTG 93072	
Qy	451	CTAGAGAAATACGGCGTTGTTGCAAGTCTCTTAATGTAATGTTTGTGTTATGATATTGCT 510	
Db	93071	TTTGGAAATATATGGGTTTAAGATTAGGACATTTGCATGTAATAATTTTGTGTTATGATGTTGCT 93012	
Qy	511	AGTGGTAGTATTCCTTTAGCACCTTATATGTGTGCTGGTGGTGGTGCAGATTATATAAG 570	
Db	93011	CACGGAATATTCCTTTAGCACCTTACGTGTGCTGGTGTATTTGGTGGGATTTATGTAANA 92952	
Qy	571	TTTTTAGGTATATCATTTGCTTAAGTTTCTTATCAAGTTTAAAGTTTGGTCTCACTACCCT 630	
Db	92951	TTTATAGGAGTTTCATTGCTTAAATTTTCTTATCAGCTTAAATTTTGGAGTAAACTACCCA 92892	
Qy	631	CTAAATGTTAATCTATGTTTGTGGGGGTTATTACCATAAGGTTGTAGGTGATAGG 690	
Db	92891	CTAAGTATTCGTACTATGCTATTTTGGTGGGTTATTCATATAAGTTATGGGTAGTAAA 92832	
Qy	691	CATGAGAGAGTAGAAATAGCTTACCCTACTGCAATTTATCTGACCTTCTAGAACTACT 750	
Db	92831	TATGATAGGGTGAAGTGGTTTATCATCTGTTTCAAGTTAAATCTGTTCTTCTAAATGACA 92772	
Qy	751	TCAGCTTCTGCTACTTTTAAATCTGATATTTTGGTTGGGAGATTGGATTAGATT 807	
Db	92771	TTTGTCTGCTAAATTTAGATATTTGATTTTGGTTGGTGGGATTAGATT 92715	
RESULT 11			
CR925678_14/c			
WPCOMMENT			

Sequence split into 16 fragments LOCUS CR925678 Accession CR925678

Fragment Name	Begin	End
CR925678_00	1	110000
CR925678_01	100001	210000
CR925678_02	200001	310000
CR925678_03	300001	410000
CR925678_04	400001	510000
CR925678_05	500001	610000
CR925678_06	600001	710000
CR925678_07	700001	810000
CR925678_08	800001	910000
CR925678_09	900001	1010000
CR925678_10	1000001	1110000
CR925678_11	1100001	1210000
CR925678_12	1200001	1310000
CR925678_13	1300001	1410000
CR925678_14	1400001	1510000
CR925678_15	1500001	1512977

Continuation (15 of 16) of CR925678 from base 1400001 (CR925678 Ehrlichia ruminantium strain 46.8%; Score 380.6; DB 1; Length 110000; Best Local Similarity 69.1%; Pred. No. 6.1e-60; Matches 578; Conservative 0; Mismatches 229; Indels 30; Gaps 3;

QY	1	ATGAATTTACAAAAGATTTGTTGTAGGTGTACGCTCAGTACATTTCTGTTTCTTTCTTTAT-- 58
Db	71320	ATGAATTTACAAAAGATTTGTTGTAGGTGTACGCTCAGTACATTTCTGTTTCTTTCTTTACTA 71261
QY	59	-CTGATGTTGCTTTTCTGATCAAAATTTTCTGAAGGAGGAGAGACTTTATATAGGT 117
Db	71260	CCTGAAGCTGTATATCTGATATGGATGTTTCAGAAAATAGAAAGTAGGTTTATGTCAGGT 71201
QY	118	AGTCAGTATAAAGTTGGTATTCCTCAATTTTAGTAATTTTTCAGCTCAAGAAACAATTCCT 177
Db	71200	GTTTCAGTATAGGACAGGTATACCAAAATTTTGATAATTTTTCAGCTTCAGAAAACAATCCCT 71141
QY	178	GGTATTACAAAAGATTTTGGTTAGGTCTTGA-----TAAGTCTGAGATAAATACT 231
Db	71140	GGTTAACTAAGGGTGTATGATTAAGATTGATTTGGATCTAAGTAATCTGATATACAAA 71081
QY	232	CACAGCAATTTTACAGCATATGACCCCTACTTATGCAAGCAGTTTTCAGGGTTTACT 291
Db	71080	CGTGCTAAATTTTCAAGATTATATAATCTCTACTTATAGTACCAGTTCTACAGGATAGGT 71021
QY	292	GGTATCATTTGGATTTATGTTATGACTTTAGGGTAGAATTTCAAGGTTCTTATGAGAT 351
Db	71020	GGTATGTTGGGTACTATTTTGATATATATAAGAAATGGAATTTTGAACAATCTTATTCAGT 70961
QY	352	TTTGAACCTGAAAGACAATGGTACCTCGAGAAATAGCAAGCTACAAAATTTTGTCTTTG 411
Db	70960	TTTCGGAATTTGAAGGCAGTGATCTCTGAAGGAGTCAAAAGTCATAGTTTGTGCTGTA 70901
QY	412	TCTCGAAATGCTA-----CAAATAGTGATATAAGTTTATAGTA 450
Db	70900	TCTCGTCAAGATAATGTCAGCTCCTAATACTGACTCAAGTAATAATGATTTGTAGTG 70841
QY	451	CTAGAGAAATACCGCGTTGTTGACAAGTCTCTTAATGTTAAAGTTTGTATGATATGCT 510
Db	70840	TTTGAAGAAATTAAGGGGTAAAGATTAGGACATTCATGTAATTTTGTATGATGTTGCT 70781
QY	511	AGTGGTAGTATTCCTTTAGCACCTTATATGTGTGCTGGTGTGGTGCAGATATATAAAG 570
Db	70780	CACGGAATATTCCTTTAGCACCTTACCGTGTGCTGTTGTTGTTGGGATTAATGTA 70721
QY	571	TTTTTAGGTATATCATTTGCTCAAGTTTCTTATCAAGTTAAGTTTGGTGTCACTACCT 630
Db	70720	TTTATAGGAGTTTCAATTCGCTAAATTTCTTATCAGCTTAAATTTTGGAGTAACTACCCA 70661
QY	631	CTAAATGTTAATACTATGTTTGTGGTGGGGTTATACCATAGTTTGTAGTGATAGG 690
Db	70660	CTAAGTATTCGTACTATGCTATTTTGTGGTGGGGTTATATCATAAAGTTATGGGTAGTAA 70601
QY	691	CATGAGAGAGTGAATAATAGCTTACCATCTCTACTGCAATTTATCTGACGCTCTCTGAACACTACT 750

Db	70600	TATGATAGGTCGAAAGTGGTTTATCATCTGTTCAAGTTAAATACCTGTTCTTAAATGACA 70541
QY	751	TCAGCTTCTGCTACTTTTAAATCTGATTTATTTGGTTGGGAGATTGGATTAGATT 807
Db	70540	TTTGTCTCTGCTAAATTTAGATATTGATTTATTTGGTTGGAGTTGGGATTAGATT 70484
RESULT 12		
AF319940/c		
LOCUS	19449 bp	DNA linear BCT 05-JUL-2005
DEFINITION	Ehrlichia ruminantium isolate Senegal map1 gene cluster, partial sequence.	
ACCESSION	AF319940	
VERSION	AF319940.4	GI:66957871
KEYWORDS	Ehrlichia ruminantium (heartwater rickettsia)	
SOURCE	Ehrlichia ruminantium	
ORGANISM	Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales; Anaplasmataceae; Ehrlichia.	
REFERENCE	1 (bases 1 to 19449)	
AUTHORS	Bekker,C.P., Bell-Sakyi,L., Paxton,E.A., Martinez,D., Bensaïd,A. and Jongejans,F.	
TITLE	Transcriptional analysis of the major antigenic protein 1 multigene family of Cowdria ruminantium	
JOURNAL	Gene 285 (1-2), 193-201 (2002)	
PUBMED	12039046	
REFERENCE	2 (bases 1 to 19449)	
AUTHORS	Bekker,C.P., Postigo,M., Taoufik,A., Bell-Sakyi,L., Ferraz,C., Martinez,D. and Jongejans,F.	
TITLE	Transcription Analysis of the Major Antigenic Protein 1 Multigene Family of Three In Vitro-Cultured Ehrlichia ruminantium Isolates	
JOURNAL	J. Bacteriol. 187 (14), 4782-4791 (2005)	
PUBMED	15995193	
REFERENCE	3 (bases 1 to 19449)	
AUTHORS	Bekker,C.P.J., Paxton,E.A., Bell-Sakyi,L., Martinez,D., Bensaïd,A. and Jongejans,F.	
TITLE	Direct Submission	
JOURNAL	Submitted (07-NOV-2000) Bacteriology, Faculty of Veterinary Medicine, Yalelaan 1, Utrecht 3584CL, The Netherlands	
REFERENCE	4 (bases 1 to 19449)	
AUTHORS	Bekker,C.P.J., Taoufik,A. and Jongejans,F.	
TITLE	Direct Submission	
JOURNAL	Submitted (09-NOV-2004) Virology, Utrecht University, Yalelaan 1, Utrecht 3584CL, The Netherlands	
REMARK	Sequence update by submitter	
COMMENT	On Jun 6, 2005 this sequence version replaced gi:55620638.	
FEATURES	Location/Qualifiers	
source	1..19449	
	/organism="Ehrlichia ruminantium"	
	/mol_type="genomic DNA"	
	/isolate="Senegal"	
	/db_xref="taxon:779"	
gene	600..1523	
	/gene="map1-14"	
CDS	600..1523	
	/gene="map1-14"	
	/note="member of multigene family"	
	/codon_start=1	
	/transl_table=11	
	/product="MAP1-14"	
	/protein_id="AAV59045.1"	
	/db_xref="GI:66957874"	
	/translation="MDNVVVFSCSLGFLSLVTLVMAHNAFSSPLPVDFSENESENVGF YTSQYSLEVFPKFSVSAVKYKHEKODKELTIFSLKBENTELKNDKGFKKGYNPVYN RNYTGSAGIYSGGGLRIELEGAFKFDVDKQKYQDNRYFALSDKEETSGQPDQ PTPVPVPVPAQPSKPTTGYNVVTAKEGLSIISLTNACYDVIIGNSOLIPSCV IGIGGINTFLGVINIKTIYKAKVGMGLLSPKTIIFINGVYVVKVPNDSFKNVSIXYQ HELEKDPGHIEPIIFNFSDYVGGEVGLRFL"	
	2400..3284	
gene	/gene="map1-13"	
	2400..3284	
CDS	/gene="map1-13"	



```
/note="member of multigene family"
/codon_start=1
/transl_table=11
/product="MAP1-13"
/protein_id="AA59046.1"
/db_xref="GI:66957875"
/translation="MSYSKSFCLMAVILLSPYVSLAVNLNENIYKGFYAGIQYKPAKY
HLSYLDKEDGYNTIGAFALKKFSKVRNDANINNTLATTASENNFTIGYDPHYENS
YLGISGALGYNYNGRVESEIISERFLKNEGKYLIDHEKYFALARSASGYGRVTRV
PSYNEVYIIMDNGRDIIFNACVDYINIHGLITYSCVGFADLVDFLGKYLK
PSYNTKIGISYVPVSSNIIIAVEGYTHGLLSRRFDKIPVNSYMQSOLNSVDTTASALL
NIRYGGSIGVRFLGSL"
3373..4083
/gene="un-1"
3373..4083
/gene="un-1"
/codon_start=1
/transl_table=11
/product="UN-1"
/protein_id="AA59047.1"
/db_xref="GI:66957876"
/translation="MNKKIKFKYLLCGLISFLSLTSFNYPNPIISNVSSCISSPFN
SIIIEHSLNATYSIDIIINSTNPLSYTLIVPNSYQLYSEBEKELVKISTNKLIVHI
PVHNYCLSNVCNVIKDDKYTKSSIFIFHQKKVLPSENFVYTKSALDHLVLLNSNI
KKFLPSKQLYKFFYKKKFWMTSDPINSNFHIYQSKKQVLSLPSKVNKNIERAISTVR
YLNLLHTFLRYNFKVIL"
4080..4907
/gene="map1-12"
4080..4907
/gene="map1-12"
/note="member of multigene family"
/codon_start=1
/transl_table=11
/product="MAP1-12"
/protein_id="AA59048.1"
/db_xref="GI:66957877"
/translation="MNYKNIFILTCILFLIPSPVSSLASNDNTLSPVQVTSVOYKPS
ISYSPKSPGIONEIVELISLKNLVNTNFIISKGQNDIKNFISYPNPTYKNSP
LGIGATGVSNRYRIELEFYEQFDLKPSEYFHKDAYKYFIKSTPSPHHLFKN
NDITVSLNVCDIIPKNTKIPVLCFGADIDVDFDKVHFVKYSYQAKIGVSYFI
LPNALFVDSGFSYHLSNKKTHIPTINIMDPPLPDSSSAKFNVLSSSGFGRIFHI"
4923..5804
/gene="map1-11"
4923..5804
/gene="map1-11"
/note="member of multigene family"
/codon_start=1
/transl_table=11
/product="MAP1-11"
/protein_id="AA59049.1"
/db_xref="GI:66957878"
/translation="MKKLYNLFILVLLATCFLPKFAFSTFHDNTIAKPITYVNYOPT
VSNFRHIESNFDTKKPIEVNINSTGISARYFKNREPAPYSSHNKHRYEHNKDLA
FTSIGSLKNFKIEAGSYKVFDFVRNAYOAGHNIIFAXPRETNSYGMPLDRPS
LXNKTGYTLKNGGIAIISNMILCYKOSNNPTPYICFICGDFIEIFDITTKIRAA
YQKIGISYPLTSRTNLLISQYHKVIGNQFKELPTLIQVELKRLPKRQPEYDVTALL
TLDIEYFSGEGLSFTL"
5832..6626
/gene="map1-10"
5832..6626
/gene="map1-10"
/note="member of multigene family"
/codon_start=1
/transl_table=11
/product="MAP1-10"
/protein_id="AA59050.1"
/db_xref="GI:66957879"
/translation="MNKNISITKVYIXITILSFILLPIQSPSALIDNIKNEEYSNVI
TTOPKPTVLYEKDLKEINASKYNNDTSKNDIIIOHDTKEHNNTSPSGSVGS
SLGRLEEGSHEKHFHONSIIISKISQYQYKAYPATIDNTYNTNNNTITLSMW
NTCYDITIGNSSTWPLYCTGIGGDIINIFNATHLRFAYQKGIGISYQLNNFFLFADT
YHKIMGNKFDLYIHNSNITPMLANIDIDYGFSGEVLRIIFNKL"
6643..7512
gene
CDS

/note="member of multigene family"
/gene="map1-9"
6643..7512
/note="member of multigene family"
/codon_start=1
/transl_table=11
/product="MAP1-9"
/protein_id="AA59051.1"
/db_xref="GI:66957880"
/translation="MKKINIILNIVILLASLAFSTKSSAVSLDNIESPQIYLGTYWK
LSISHTNFSVAXNQTVDIIGLKNVNTEDDIIKTAKNFNTQVNPPIKNNFTGSGG
ALGYXSGRLRELEASQYODPVKKSKYKTNDAHRYFALVRNKDDKAFQPODDVLYR
KRYRSPYSFMRNDGISIASWMENGCYDLPSNFKISSYACIGIGGDFIEFEAMKVKF
AYQAKLGISYIISPSVNLFIADIIYHKSVGNQFNLRVQVYHRLRTPITTSALAKLNI
GYFGGEVGRFIF"
7537..8385
/gene="map1-8"
7537..8385
/gene="map1-8"
/note="member of multigene family"
/codon_start=1
/transl_table=11
/product="MAP1-8"
/protein_id="AA59052.1"
/db_xref="GI:66957881"
/translation="MNKNKSIIGTALTFLAFSPIESFSANQTDITITAPISGIY
FTQYKPGISNFSNFSAKETNFNTQKLVRLKDKAKESNLLGTTNFEDTYSVKFQNNI
ISFGIIGYATSKIRLEIEGAYESPVKSPVGYSKDNSYVYFALARSMTKDNPKF
TVLKNGLSVASIMINGCYDFALDPSYICAGIGEDFIEFFDALHIKIAYOQKLG
ISYRV9PKISLFDVGYHHIIGNQPKNLVHHAVELSEFPKSSAVATLDDIGYLGGEV
GARFIF"
8396..9247
/gene="map1-7"
8396..9247
/gene="map1-7"
/note="member of multigene family"
/codon_start=1
/transl_table=11
/product="MAP1-7"
/protein_id="AA59053.1"
/db_xref="GI:66957882"
/translation="MGKFNKYNTLLIGIMLALLPTKSFSSESINNDNSONFKLYIS
GOYKPGVFRNPNFSKATNVTKQLGLSGTSRDKDKFNSNTYIPDQDSSTGFG
AAVGYMSYKIRFELESSEYEFQVNIHDCETGTEPCRYFALARSVGRGWPEQNNYT
VMKNTGLSMTSVFNGCYNILSNKLEISPVCGICIGGDFIDFFDATHIKLAYQKLG
SYSLLPDVTLLFVDGYVHVRVNNQFNKLVLPVQLKYEKPKITSAVATMDVTFGGEVG
IRPFPNS"
9267..10163
/gene="map1-6"
9267..10163
/gene="map1-6"
/note="member of multigene family"
/codon_start=1
/transl_table=11
/product="MAP1-6"
/protein_id="AA59054.1"
/db_xref="GI:66957883"
/translation="MNKNKLIILNTALVFLSLFLPQALSIPTNNSISKYSIGIYFSG
SYKLEFPSLDNFSIKETTFNTQVIGLSKKKANVKDILGYHVAFNEPYTPTFQDTVS
GFGSTGVISYNNKLRSEIEVSYEKFDAGIPEGSIYGDYIEDAYRYFGLARETSYLSGS
SLTTPKHYHYVMKNNGVSIITSLMVNNCYQFSTQSNKILPYICGSGVGLDLVOFFNK
LHIKACQVKLGTSYLSPHYQLFANIYYHEVGNFNKLNVPVILPKDITSTTELGNN
VSATSTLNIYSFGSEIGLAFIL"
10172..10795
gene
CDS

Query Match 45.0%; Score 365.6; DB 1; Length 19449;
Best Local Similarity 68.8%; Prod. No. 5.8e-57;
Matches 541; Conservative 0; Mismatches 224; Indels 21; Gaps 2;

Qy 43 TTGTGTTTTTCTATGATGCTGCTTTTCTGATGCAAAATTTTCTGAAGGAGGAGA 102
Db 19427 TTGTTTCTTTTACTACTGAACGTGCTATATCTGATGCTGATATTTTCAGAAATAGAAGT 19368
```

```
QY 103 GGACTTTATATAGGTAGTCAGTATAAAGTTGGTATTCCTCCAAATTTAGTAAATTTTCAGCT 162
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 19367 AGGTTTATGCGAGGTGTCATATATAGCAGGTATACCAATTTTGTATATTTTCAGCT 19308
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 163 GAAGAAACAATTCCTCGTATTTACAAAAGAGATTTTTCGGTTAGGCTTCA-----TAAG 216
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 217 TCTGAGATAAATCTCACAGCAATTTTACAGCATATATGACCCCTACTTATGCAAGCAGT 276
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 19247 TCTGGTATTACAAAACGTGCTAATTTTACAGGATTTGTAATCTCTACTTTATAGTACTAGT 19188
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 277 TTTGCAAGGTTTATGCTGATCATTTGATATATGTTAAATGACTTTAGGGTAGAATTTGAA 336
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 19187 TCTACAGGATAGGTGATGTTTGGGTACTATTTTGTAAATATAGAAATGAGTTTGA 19128
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 337 GGTCTTATGAGAAATTTTGAACCTGAAAGCAATGTTTACCCCTGAGAAATGCCAAAGCTAC 396
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 19127 ACATCTTATCAAAATTTCCGAGTTGAAAGGCGAGTGGTATCTCGAAGGAGTCAAAGTCAT 19068
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 397 AAATTTTTCCTTGTCTCGAAATGC-----TACAAATAGTGAATAAG 441
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 19067 AAGTTCTGTGCTGTATCTCGTAATGCAACTCCTPAGTACTGATTCAGAGTAATAATAGT 19008
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 442 TTTATAGTACTAGAGAAATACGGCGTTGTTGACAGTCTCTTAATGTAATGTTTCTTAT 501
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 19007 TTTGTAGTCTAGAAATATAGGGTTAAGGTTAGGACATGAAATGTAATTTTGTAT 18948
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 502 GATATTGCTAGTGTAGTATTCCTTTAGCACCTTATATGTGCTGCTGTTGTCAGAT 561
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 18947 GATGTTGCTCACGAAATATTCCTTTAGCACCTTATGTGTGCTGCTGTTTGGGGAT 18888
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 562 TATATAAGTTTATAGTATATCATTTGCCCTAAGTTTCTTATCAAGTTAAGTTTGTGTC 621
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 18887 TATGTAATATTTATAGGAGTTTCATTTGCCCTAAATTTTCTTATCAGCTTAAATTTGGAGTA 18828
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 622 AACTACCCTCTAAATGTTAATACTATGTTGTTGGTGGGGTTATTACCATTAAGTTGTA 681
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 18827 AACTATCCACTAGTATTCCTACTATGCTATTTGTTGGTGGGGTTATTATCAAAAGTTATA 18768
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 682 GGTGATAGGCAATGAGAGAGTAAATAGCTTACCATCTCTACTGCTCATTTATCTGACGTTCTT 741
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 18767 GGTAGTAAATATGATAGGGTGAAGAGTAGTTTATCATCTGTTTCAAGTAACTGTTCTT 18708
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 742 AGAATCTTTCAGCTTCTGCTACTTTAAATACTGATTTATTTGGTGGGAGTTGATTT 801
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 18707 AANAATGACATTTGTTTCTGCTAATTTAGATATTGATTTATTTGGTGTGGAAGTTGGGATT 18648
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 802 AGATTT 807
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 18647 AGATTT 18642
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 13
AR303095 LOCUS 840 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 23 from patent US 6544517.
ACCESSION AR303095
VERSION AR303095.1 GI:31691735
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 840)
AUTHORS Rikihisa,Y. and Ohashi,N.
TITLE Outer membrane protein of Ehrlichia canis and Ehrlichia chaffeensis
JOURNAL Patent: US 6544517-A 23 08-APR-2003;
The Ohio State University Research Foundation; Columbus, OH
FEATURES
source Location/Qualifiers
1..840
/mol_type="genomic DNA"
ORIGIN
1 (bases 1 to 840)
```

```
Query Match 20.0%; Score 162.2; DB 6; Length 840;
Best Local Similarity 53.8%; Pred. No. 1.3e-19;
Matches 432; Conservative 0; Mismatches 353; Indels 18; Gaps 4;
QY 22 GTAGGTGTTTACGCTGAGTACATTTGTTTCTTATCTGATGGTCTTTTCTGATGCA 81
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 34 GTACTTGCATCTCTATTATCATTTTATCTATTGAATCCTTTTCAGCTATAAATCATAAT 93
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 82 AATTTTCTGAAGGAGGAGAGAGACTTTATATAGTAGTCAGATATAAAGTTGGTATTTCCC 141
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 94 CATACAGAAATAAACACTAGTGGTATATATTTACAGGCGAGTATAGACCAGAGTATCC 153
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 142 AATTTTATAGTAATTTTTCAGCTGAAGAAACAATTCCTGCTATTACAAAAAGATTTTGG 201
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 154 CATTTTACAAATTTCTCAGTAAAGAAACTAATGTTGATACAACTAGTAGATAT 213
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 202 TTAGGTCTTGATAAGTCT---GAGATAAATACTCACAGCAATTTTACACGATCATATGAC 258
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 214 AAAAAGAGTGGCTCTTCTATCGATCCTTAACACTTATTCAAAGTCCATATACT 273
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 259 CCTACTATGCAAGCAGGTTTTGACAGGTTTGTAGTGTATCATTTGGATATATGTTAATGAC 318
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 274 GTTACATTTCAAGATAAATGCTGTAGTTTTCAGTGGAGCAATTTGGATATTTCTTACC 333
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 319 TTT---AGGGTAGAATTTGAAAGTCTTATGAGAAATTTTGAACCTGAAAGACAATGGTAC 375
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 334 AGTCTAAGACTTGAACCTTGAAGGTTCTTACGAAATTTTGATGTCAAAGATCCTAAAGAC 393
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 376 CTTGAGAATAGCCAAAGCTACAAATTTTGTCTGCTCGAAATGCTACAAATAGTAT 435
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 394 TACTCAGCAAAAGATGCTTTTAGGTTTTTGTCTAGCACGTAATACGTTCTACTCTGTT 453
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 436 AAT-----AAGTTTATAGTACTAGAGAAATAACGGGTTGTTGACAAGTCTCTTAAT 486
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 454 CCTGATGCTCAAAAATATACAGTTATGAAGAAATAATGGCTTATCTGTTGCAATCAATG 513
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 487 GTAAATGTTTGTATGATATTGCTAGTGTAGTATTTCTTTAGCACCTTATATATGTTGCT 546
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 514 ATCAATGTTTGTATGATCTATCTTTTAAATTTAGTGTATCATCTATATATATGCA 573
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 547 GGTGTTGTCAGAGATATATATAAGTTTTTGTAGTATATCATTTGCCCTAAGTTTTCTTCAA 606
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 574 GGTATTGTTGAGAGATTTTCATTGAAATTTTGTATCTTTGCAACATTAACCTGCTTATCAA 633
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 607 GTTAAGTTTGGTGTCAACTACCCCTCTAAATGTTTAAATATATATGTTGGTGGGGTTAT 666
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 634 GGAATAACTAGGTATTAGTTATTACTTCTTCTTAAAGATTAATGTTATTTGCTGGGTAC 693
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 667 TACCATAAGGTTGTAGGTGATAGGCAATGAGAGAGTAGAAATAGCTTACCATCTACTGCA 726
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 694 TATCATAGAGTTATAGGGAATAAATTTAAATTT---AATGTTAAACCAAGTTGTACA 750
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 727 TTATCTGAGCTTCTTAGAAGTACTTTCAGCTTCTGCTACTTTTAAATACTGATTTATTTGT 786
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 751 CTTGATGAATTTCTTAAGCAACTTCTGAGTAGTACACTTAATGTTGCTTATTTTGCT 810
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 787 TGGGAGATTGGATTTAGATTTGC 809
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 811 GGTGAAGCTGGAGTAAAGTTTAC 833
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 14
AR655791 LOCUS 840 bp DNA linear PAT 13-JUN-2005
DEFINITION Sequence 23 from patent US 6893640.
ACCESSION AR655791
VERSION AR655791.1 GI:67588223
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 840)
```



```

gene
CDS
CKKDSIESTDNGNHNVMKNEGFRVISLTFNACYDMIVSNSLSVPSACIGICQGITN
FLGCTNIHTLFAKALGLGLFSLSPKTVIFANGYVVKTKDMSFTNLSVQVPLELKEAPKH
IDPIACFNADNYGGEVGLRFL"
2498. .3352
/gene="omp-1n"
2498. .3352
/gene="omp-1n"
/notes="Omp-1 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein OMP-1n"
/protein_id="AAK28659.1"
/db_xref="GI:13511830"
/translacion="MSKRSNRKFLVWMLLFTPHISLASVLNDHSMVYGIQKPAR
OHLKLLIKESANTVEVFLKQDLNLLTGIKONTNFIKYNPVYNNRNLGRSGIF
GYNNKPNRIESELVSTETHIKNGYKRIDCEKHPALAKEISGGSNPNANKYVILIN
NGLSLTSALINVCYDDEGLKHNITISYCLGFGVDITDLSKYTTTPSYOGKLGSYTV
SPQVSVEIFEGYHGLFGKFKPIPVNYPDCYFSPTPNSKPHVHTTALAMLSIGYGY
SIGIKFPL"
3459. .3905
/codon_start=1
/transl_table=11
/product="unknown function UN2"
/protein_id="AAK28660.1"
/db_xref="GI:13511831"
/translacion="MNKKSRGOVLLFAAYSIFKYINSCTSKFRLSVKCFVYNPLKYF
YKSTPLKLGKFWNYNYKTSNLYHSLFKLGIIRTTNFKISVLNFTNPRVLIIY
YKLNKKNNSNLYLNKTYINTATISFINSVILGIYGFVNYNT"
3915. .4808
/gene="omp-1q"
3915. .4808
/gene="omp-1q"
/notes="Omp-1 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein OMP-1q"
/protein_id="AAK28661.1"
/db_xref="GI:13511832"
/translacion="MSYAKVFILICLLVPSLSPALVNNDFLKNIGHFIYGGYKPK
GVPRNRPVLTNNIRBELMSSDEECSTIPHVMQSVQAQGLPPEALEELAKGLHGGY
LFTLPNTYNNIKNLGAGGVIGYSTHFRVEAEYERFNLTAPAGYLHKNFYBFA
LATMDTKPHQSAEDKYMYMNTGITLSPIINACYDFILKTRNVAEPLCLGVGGN
FIDLQVDFKFAQYQAKVGISYFVSPNIAFFIDGSPHGLNNQFSDLPVVDYSSSGFP
TTSKANKANPLTSSIGIRFIS"
4825. .5682
/gene="omp-1p"
4825. .5682
/gene="omp-1p"
/notes="Omp-1 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein OMP-1p"
/protein_id="AAK28662.1"
/db_xref="GI:13511833"
/translacion="MOKLYISFILSGLLPPKYVFCMHQNNIDGSYVTKYQLTPPH
KNFYIKEDFDQEPDGLAKITANTKFTDLKENPSPSLHOTDSYKSYONDLLIGL
SVGLFVKSRIPEEGAKYNTKRLARYSKDGYKYFAIPRKEHGFELNDRGTGYTAK
NNGISLIISNINILCSYTKAKFTPYICIGVGGDFIIFDVMRIKFAQYQKGVSPIT
SKULISNGYHKVIGNKFFELLPVYQFVELKRLVNTNKTSKDIDQDVTASLTNLNLEHFS
SEIGLSFIF"
5695. .6513
/gene="omp-1t"
5695. .6513
/gene="omp-1t"
/notes="Omp-1 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein OMP-1t"
/protein_id="AAK28663.1"
/db_xref="GI:13511834"
/translacion="MTWNNKHYCYIVTVITLFFLLPIESLSALIGNVKDLKVSS
TVVSSQYKPSIIFHRNFSIQEBSPKKSEBEFKIKANLNNILKSNAYNIQFDNTTFS
SGTIGYFSKGLRLEARGCYQEFNVKNNSNLSIISNNKYHSRIHDENYAITTNKLSIA

```

```

gene
CDS
STWNTCYDISNNTSIVPYLCTGIGEDLVGLFNTIHFKLAYQGVKMSYLIINNILL
FSDIYHVKNRPNKLYMQVADPNISEETIPIILAKLDIGYFGSEIGIRPMFN"
6529. .7416
/gene="omp-1u"
6529. .7416
/gene="omp-1u"
/notes="Omp-1 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein OMP-1u"
/protein_id="AAK28664.1"
/db_xref="GI:13511835"
/translacion="MTKKFNFMVNVILTELLFLPLKSFPTTYANNNTIITQKVLGISGO
YKPSIPHPKNSVEENDKVVDLIGLTDVTVITEHILRDNTKFNTHYIAKPNNEINF
SSAIGVYSCQSPRLIESSYGDVNVVKNYAVODVRYFALVREKNSNPSKPHET
SQPSDSNPKKSFYTLMKNGVVFVASVIINGCYDFSNNNTIISPYVICIGVGGDFIEFFE
VWHIKFACQSKVGISYPLSPISITIPADAHYHKVINNNKFNHNLHVKYSYELKNSPTIISA
TAKLNIEYFGGEVGMRFIF"
7438. .8277
/gene="omp-1v"
7438. .8277
/gene="omp-1v"
/notes="Omp-1 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein OMP-1v"
/protein_id="AAK28665.1"
/db_xref="GI:13511836"
/translacion="MSKKKFTITGTVALSLSLFSLIESBSAINHHHTGNTSGIYITG
QYRPGVSHFSNFSKETNVDITQLVGYKKSASSIDPNTSYNFQGPYTVTFDQNAASF
GAIGYSYPSERLLEGSYKFKDPKDYSAKADPRFALARNTSTVPDAQKYTMV
KNGLSVASIMINGCYDLSFNVLVSPVICAGIGDEFEFFDTLHILKAYOGKLGISY
YFPKINVPAAGYHRYVIGNKFKNLVNHNVTLDPEFKATSAVATLNVAYFGGEAGVK
FTF"
8300. .8890
/codon_start=1
/transl_table=11
/product="unknown function UN3"
/protein_id="AAK28666.1"
/db_xref="GI:13511837"
/translacion="MIYKELTRVGEYILAYLSFILSYIVFLVLNIIRYNSLAICVI
SLLRNIENVTSTKLIKDKCRDTRFNNNLYLQKPLNLQIFYGIFSRFNQNTLI
IPNDSKCFGYTLMNPAHYTYTLTGEYRNFIDILYENIICQKLLINVRNSVLNQ
HNKNTLVIIPINAREFSINVRNISINKSSYBC"
8880. .9731
/gene="omp-1w"
8880. .9731
/gene="omp-1w"
/notes="Omp-1 family member"
/codon_start=1
/transl_table=11
/product="major outer membrane protein OMP-1w"
/protein_id="AAK28667.1"
/db_xref="GI:13511838"

```

```

Query Match      20.0%; Score 162.2; DB 1; Length 27190;
Best Local Similarity 53.8%; Pred. No. 4e-20;
Matches 432; Conservative 0; Mismatches 353; Indels 18; Gaps 4;

Qy  22  GTAGGTGTACGCTAGTACATTTCTTTTTTTTCTTATCTGATGGTCTTTTCTGATGCA 81
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  7471 GTACTTGCATCTCTATTATCATCTTATCTATTGAATCCTTTTCAGCTATAAATCATAAT 7530
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy  82  AATTTTCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 141
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  7531 CATACAGGAAATAACACTAGTGGTATATATATACAGGCGAGTAGATAGCAGGAGTATCC 7590
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy  142  AATTTTAGTAATTTTTCAGCTGAAGAAACAATCTCTGGTATTACAAAAAGATTTTTCGCG 201
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  7591 CATTTTAGCAATTTCTAGTAAAGAAACTAATGTTGATACANTACACTAGTAGATAT 7650
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy  202  TTAGGTCTTGTAAGTCT---GAGATAAATACTCAGACAAATTTTACACGATCATATGAC 258
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 7651 AAAAAAGTGGCTCTTATCGATCCTAACACATTTATTCAAACTTTCAAGGTCCATATACT 7710  
Qy 259 CCTACTTATGCAAGCAGTTTTCAGGGTTTAGTGGTATCAATGCGATATTAATGTTAATGAC 318  
Db 7711 GTTACATTTCAAGATAATGCTGCTAGTTTCAGTGGAGCAATTGGATATCTTACCCGAA 7770  
Qy 319 TTT---AGGGTAGAATTTGAAGTTCCTATGAGAAATTTGAACTCTGAAGCAATGGTAC 375  
Db 7771 AGCTAAGACTTGAACCTGAAGTTCTTACGAAAAATTTGATGTCAAAGATCCTAAAGAC 7830  
Qy 376 CCTGAGAATAGCCAAAGCTACAAATTTTTCCTTGTCTCGAAATGCTACAAATAGTGAT 435  
Db 7831 TACTCAGCAAAAGATGCTTTTAGGTTTTTGTCTAGCAGTAATACGCTCTACTACTGTT 7890  
Qy 436 AAT-----AAGTTTATAGTACTAGAGAATAACGGCGTTGTTGACAAAGTCTCTTAAT 486  
Db 7891 CCTGATGCTCAAAAATATACAGTTATGAAGAATAATGGCTTATCTGTTGCATCAATCATG 7950  
Qy 487 GTAATGTTGTTTATGATATTGCTAGTGGTAGTATTCCTTTAGCACCTTATATGTTGCT 546  
Db 7951 ATCAATGGTTGTTATGATCTATCTTTTAATAAATTTAGTCGPATCACCTTATATATGCA 8010  
Qy 547 GGTGTTGGTGAGATTATATAAAGTTTTTAGGTATATCATGCTTAAGTTTTCTTATCAA 606  
Db 8011 GGTATTTGGTGAAGATTTTCATTTGAATTTTTTGATACTTTGCACATTAACCTTGCTTATCAA 8070  
Qy 607 GTTAAGTTTGGTGCAACTACCTCTAAATGTTAAATGTTTGTGTTGGTGGGGTTAT 666  
Db 8071 GGAAAACTAGGTATTAGTTATTACTTCTTCCCTAAGATTAAATGTTTGTGTTGGGTAC 8130  
Qy 667 TACCATAAGTTGTAGGTGATAGGCATGAGAGAGTAGAATAAGCTTACCCTACTGCA 726  
Db 8131 TATCATAGATTATAGGGAATAAAATTTTAAAAATTT---AAATGTTAACCATGTTGTTACA 8187  
Qy 727 TTATCTGACGTTTCTAGAACTACTTTCAGCTTCTGCTACTTTTAAATACCTGATTATTTGGT 786  
Db 8188 CTTGATGAATTTCTTAAGCAACTTCTGAGTAGCTACACTTAATGTTGCTTATTTTGGT 8247  
Qy 787 TGGGAGATTGGATTTAGATTTGC 809  
Db 8248 GGTGAAGCTGGAGTAAAGTTTAC 8270

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: December 28, 2005, 10:46:11 ; Search time 620 Seconds  
(without alignments)  
8739.347 Million cell updates/sec

Title: US-10-731-554-45  
Perfect score: 813  
Sequence: 1 atgaattacaagaattgtg.....ttggatttagattggcgcta 813

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues  
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_21.\*  
1: Geneseq1980s.\*  
2: Geneseq1990s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001as.\*  
5: Geneseq2001bs.\*  
6: Geneseq2002as.\*  
7: Geneseq2002bs.\*  
8: Geneseq2003as.\*  
9: Geneseq2003bs.\*  
10: Geneseq2003cs.\*  
11: Geneseq2003ds.\*  
12: Geneseq2004as.\*  
13: Geneseq2004bs.\*  
14: Geneseq2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	813	100.0	813	6	ABK68878 DNA encod
2	162.2	20.0	840	6	AAX34754 DNA encod
3	162.2	20.0	840	6	ABK63287 DNA encod
4	162.2	20.0	840	9	ADA09756 E. chaffe
5	162.2	20.0	840	14	ADW04249 Ehrlichia
6	150.8	18.5	852	2	AAX34755 DNA encod
7	150.8	18.5	852	6	ABK63288 DNA encod
8	150.8	18.5	852	9	ADA09758 E. chaffe
9	150.8	18.5	852	14	ADW04251 Ehrlichia
10	147.2	18.1	840	6	ABK68876 DNA encod
11	147.2	18.1	843	6	ABK63299 DNA encod
12	147.2	18.1	843	9	ADA09780 E. canis
13	147.2	18.1	843	14	ADW04273 Ehrlichia
14	146.6	18.0	726	6	ABK63306 DNA encod
15	146.6	18.0	726	9	ADA09794 E. canis
16	146.6	18.0	726	14	ADW04287 Ehrlichia
17	146.4	18.0	849	6	ABT11835 E chaffe
18	146.4	18.0	849	14	ABK63350 Ehrlichia
19	135	16.6	828	2	AAX34756 DNA encod

20	135	16.6	828	6	ABK63289 DNA encod
21	135	16.6	828	9	ADA09760 E. chaffe
22	135	16.6	828	14	ADW04253 Ehrlichia
23	133.8	16.5	840	6	ABK63305 DNA encod
24	133.8	16.5	840	9	ADA09792 E. canis
25	133.8	16.5	840	14	ADW04285 Ehrlichia
26	129.6	15.9	843	2	AAX34745 DNA encod
27	129.6	15.9	843	6	ABK63278 DNA encod
28	129.6	15.9	843	9	ADA09738 E. chaffe
29	129.6	15.9	843	14	ADW04231 Ehrlichia
30	126.4	15.5	837	2	AAX34747 DNA encod
31	126.4	15.5	837	6	ABK63280 DNA encod
32	126.4	15.5	837	9	ADA09742 E. chaffe
33	126.4	15.5	837	14	ADW04235 Ehrlichia
34	126.4	15.5	4683	4	AAS07578 DNA encod
35	126.4	15.5	4683	3	AAC68716 Ehrlichia
36	124.8	15.4	4683	2	AAV07179 Ehrlichia
37	123.2	15.2	4683	2	AAV07179 Ehrlichia
38	122.4	15.1	867	2	AAX34759 DNA encod
39	122.4	15.1	867	6	ABK63291 DNA encod
40	122.4	15.1	867	9	ADA09764 E. canis
41	122.4	15.1	867	14	ADW04257 Ehrlichia
42	121.8	15.0	843	2	AAX34748 DNA encod
43	121.8	15.0	843	3	AAC68705 Ehrlichia
44	121.8	15.0	843	6	ABK63281 DNA encod
45	121.8	15.0	843	9	ADA09744 E. chaffe

ALIGNMENTS

RESULT 1  
ABK68878  
ID ABK68878 standard; DNA; 813 BP.

AC ABK68878;

XX 02-JUL-2002 (first entry)

XX DNA encoding Ehrlichia canis p28-9.

XX Ehrlichia canis infection; vaccine; serodiagnostic; gene; p28; ss;  
antibacterial.

XX Ehrlichia canis.

XX WO200222782-A2.

XX 21-MAR-2002.

XX 12-SEP-2001; 2001WO-US028759.

XX 12-SEP-2000; 2000US-00660587.

XX (RERE-) RES DEV FOUND.

XX Walker DH, Yu X, McBride JW;

XX WPI; 2002-351882/38.

XX P-PSDB; AAU96118.

XX New recombinant homologous 28 kilodalton immunodominant protein from  
Ehrlichia canis, useful for treating Ehrlichia canis infections.

XX Claim 5; Fig 16; 106pp; English.

XX The invention relates to a recombinant homologous 28 kDa immunodominant  
protein, p28 (I) of Ehrlichia canis. (i), a 28-kDa antigen preferably  
dispersed in a pharmaceutically acceptable carrier, is useful for  
inhibiting E. canis infection in a subject. (i) is useful in the  
development of vaccines and serodiagnostics that are particularly  
effective for disease prevention and serodiagnosis. ABK68878-ABK68878  
represent the 28-kDa antigen coding sequences and PCR primers of the







22 GTAGGTGTTACGCTGAGTACATTTGTTTTTCTTATCTGATGCTGCTTTTCTCTGATGCA 81  
34 GTACTTGCATCTCTATTATCATCTTATCTATTTGAATCTTTTCCAGCTATPAAATCATAA 93  
82 AATTTTTCTGAAGGGAGGAGGAGCTTTATATATAGTGTAGTACAGTATAAAGTTGGTATTC 141  
94 CATACAGAAATTAACACTAGTGTATATATATTA CAGGGCAGTATAGACCAGGAGTATCC 153  
142 AATTTTAGTAAATTTTTCAGCTGAAGAAACAAATCTCTCGGTATTACAAAAGAGATTTTTTGGC 201  
154 CATTTTAGCAATTTCTCAGTAAAGAAACAAATAGTTGATACAAATCAACTAGTAGAGTAT 213  
202 TTAGGTCCTTGATAAGTCT---GAGATAAATACTCA CAGCAATTTTTTACAGCATCATATGAC 258  
214 AAAAAAGTGGCTTCTATCGATCCTAACACTATTTCAAACTTTTCAAGGTCCATATACT 273  
259 CCTACTTATGCAAGCAGTTTTTGCAGGGTTTTAGTGGTATCATATGGATATTAATGTTAAATGAC 318  
274 GTTACATTTCAAGATAAATGCTGCTAGTTTTCAGTGGAGCAATTTGGATATCTTCAACCCGAA 333  
319 TTT---AGGGTAGAATTTGAAGTCTTATGAGAAATTTTGAACTCTGAAAGACAATGGTAC 375  
334 AGTCTAAGACTTTGAACCTTGAAGGTTCTTACGAAAAAATTTTGATGTCAAGAGATCCTAAAGAC 393  
376 CCTGAGAATAGCCAAAGCTACAAATTTTTTTTGTGTTTCTCGAATGCTACAAATAGTGTAT 435

DR WPI; 2005-0

DR WPI; 2005-0

DR P-PSDB; ADW04250.  
 XX New polynucleotide encoding an outer membrane protein (OMP) of Ehrlichia  
 PT canis or Ehrlichia chaffeensis, useful in preparing a composition for  
 PT diagnosing or preventing E. canis or E. chaffeensis infection.  
 XX  
 XX Claim 7; SEQ ID NO 23; 122pp; English.  
 XX  
 CC The invention relates to nucleic acid sequences encoding outer membrane  
 CC proteins (OMP) of Ehrlichia chaffeensis (designated as OMP proteins) and  
 CC Ehrlichia canis (designated as P30F proteins). The OMP polynucleotide is  
 CC useful in preparing a composition for diagnosing, treating or preventing  
 CC an infection with Ehrlichia canis or Ehrlichia chaffeensis. The present  
 CC sequence is the Ehrlichia chaffeensis OMP DNA.  
 XX  
 XX Sequence 840 BP; 279 A; 128 C; 136 G; 297 T; 0 U; 0 Other;  
 SQ

Query Match 20.0%; Score 162.2; DB 14; Length 840;  
 Best Local Similarity 53.8%; Pred. No. 2.1e-28;  
 Matches 432; Conservative 0; Mismatches 353; Indels 18; Gaps 4;

QY 22 GTAGGTGTTAGCTGAGTACATTTGTTTTTTTCTTATCTGATGCTGCTTTTCTGATGCA 81  
 DB 34 GTACTTGCATCTCTATTATCATCTTATCTTATGATCTTTTTCAGCTATAAATCATAAT 93  
 QY 82 AATTTTCTGAAGGAGGAGGAGGACTTTATATAGGTAGTACGATATAAAGTTGGTATCCC 141  
 DB 94 CATACAGGAAATAACACTAGTGGTATATATATACAGGGCAGTATAGCAGGAGTATCC 153  
 QY 142 AATTTTGTATTTTTCAGCTGAAGAACAACTTCTGTTATTAACAAGATTTTTCG 201  
 DB 154 CATTTTGTAGCAATTTCTCAGTAAAGAACTAATGTTGATACATACACTAGTAGGATAT 213  
 QY 202 TTAGGTCTTGAATGCTCT--GAGATAAATACTCACAGCAATTTTACACGATCATATGAC 258  
 DB 214 AAAAAAGTGGTCTCTATCGATCCTAACACTTATTCAAATTTCAAGGTCATATACT 273  
 QY 259 CCTACTTATGCAAGCAGTTTTCAGGGTTTATGAGTATCATGATGATATATGTTAATGAC 318  
 DB 274 GTTACATTTCAAGATAATGCTGCTAGTTTTCAGTGGAGCAATGGAATTTCTTACCCGAA 333  
 QY 319 TTT--AGGTTAGATTTGAGGTTCTTATGAGATTTTGAACCTGGAAGCAATGTGAC 375  
 DB 334 AGTCTAAGACTTGAATTTGAGGTTCTTACGAAAAATTTGATGCAAGATCCTTAAGAC 393  
 QY 376 CTGAGATAGCCAAAGCTACAAATTTTTCGTTTCTCGAAATGCTACAAATAGTGAT 435  
 DB 394 TACTCAGCAAAAGATGCTTTTAGGTTTTTCTGCTAGCACGTAATACGCTACTACTGTT 453  
 QY 436 AAT-----AAGTTTATAGTACTAGAGAAATAACGGCGTTGTTGACAAAGTCTCTTAAT 486  
 DB 454 CTGATGCTCAAAAATATACAGTTATGAAGAATAATGCTTATCTGTTGCATCAATCATG 513  
 QY 487 GTAAATGTTGTTATGATTTGCTAGTGGTAGTATCTTTTAGCACCTTATATCTGTGCT 546  
 DB 514 ATCAATGTTGTTATGATCTATCTTTTAAATAATTTAGTCGATCACCTTATATATGTC 573  
 QY 547 GGTGTTGTGCAGATTATATAAGTTTTTGTAGTATATCATTTGCGCTTATCTTATCA 606  
 DB 574 GGTATGTTGAGATTTTCAATGAAATTTTGTATCTTTGACATTTAACTTGTCTTATCA 633  
 QY 607 GTTAAATTTGTTGCTCAACTACCTCTAAATGTTAATACTATGTTGTTGGGGGGTTAT 666  
 DB 634 GGAATACTAGGTATTTAGTTATTTACTTCTTCTAAGATTAATGTTATTTGCTGGGTAC 693  
 QY 667 TACCATAGTTGTTAGGTAGTATGAGCATGAGAGGTAGAAATAGCTTACCTCTACTGCA 726  
 DB 694 TATCATAGATTTATAGGGAATAAATTTTAAAAATTT--AAATGTTTAAACCATGTTGTACA 750  
 QY 727 TTATCTGAGCTTCTCTAGAACTACTTTCAGCTTCTGCTACTCTTAAATACTGATTTATTTGGT 786  
 DB 751 CTTGATGAATTTCTTAAAGCAACTTCTGAGTAGCTACATTAATGTTGCTTATTTGGT 810

QY 787 TCGGAGATTGGATTAGATTTC 809  
 DB 811 GGTGAAGCTGGAGTAAAGTTTAC 833

RESULT 6  
 AAX34755  
 ID AAX34755 standard; DNA; 852 BP.  
 XX  
 XX AAX34755;  
 AC  
 XX 27-AUG-2003 (revised)  
 DT 05-JUL-1999 (first entry)  
 XX  
 XX DNA encoding OMP-1w protein.  
 DE  
 XX Outer membrane protein; OMP; Ehrlichia chaffeensis; E. canis; P30;  
 KW detection; dog; ss.  
 XX  
 XX Ehrlichia chaffeensis.  
 OS  
 XX WO9913720-A1.  
 PN  
 XX 25-MAR-1999.  
 PD  
 XX  
 PF 18-SEP-1998; 98WO-US019600.  
 XX  
 XX 19-SEP-1997; 97US-0059353P.  
 PR  
 XX (OHIS ) UNIV OHIO STATE.  
 PA  
 XX Rikihisa Y, Ohashi N;  
 PI  
 XX WPI; 1999-254290/21.  
 DR  
 XX P-PSDB; AAY06955.  
 PT Novel outer membrane proteins from Ehrlichia chaffeensis and Ehrlichia  
 PT canis.  
 XX  
 XX Disclosure; Fig 15A; 55pp; English.  
 XX  
 CC The invention provides isolated outer membrane proteins (OMP) from  
 CC Ehrlichia chaffeensis and E. canis. The E. chaffeensis proteins form part  
 CC of the OMP family and consist of proteins OMP-1, -11B to Z) shown in  
 CC AAY06943-958. The E. canis proteins form part of the P30 family and  
 CC consist of proteins shown in AAY06959-970. The proteins and genes are  
 CC used to detect E. chaffeensis in patients and E. canis in dogs. (Updated  
 CC on 27-AUG-2003 to correct OS field.)  
 XX  
 XX Sequence 852 BP; 295 A; 115 C; 141 G; 301 T; 0 U; 0 Other;  
 SQ

Query Match 18.5%; Score 150.8; DB 2; Length 852;  
 Best Local Similarity 54.6%; Pred. No. 1e-25;  
 Matches 402; Conservative 0; Mismatches 307; Indels 27; Gaps 4;

QY 96 GAGAGAGGACTTTATATAGGTAGTCGATATAAGTTGGTATTCCTCCCAATTTTAGTAATTT 155  
 DB 111 GTGCACCTGGGCTATATGTCAGTGGCAATATAAACCTACTGTTTCTCATTCTAGTAATTT 170  
 QY 156 TTCAGCTGAGAGAAACAATTCCTGTTATTAACAAAAAGATTTTTCGCTTAGTCTTGATAA 215  
 DB 171 TTCACCTTAAAGAAACCTTATCTGACACTAAAGAGTTATTTAGGACTACCAAGATATTA 230  
 QY 216 GTCT-----GAGATAAATACTCACAGCAATTTTACACGATCATATACCCCTACTTATGC 269  
 DB 231 GTCTTATACAGATATAACAACAATAAATAATCAACATTTCTTATAACACAAAAATTTCA 290  
 QY 270 AAGCAGTTTTCAGGGTTTATAGTGTATCATTTGATATTTATGTTAATGACATTT---AGGTT 326  
 DB 291 AGATAATGCTGTTAGCTTTCAGTGCAGCTGTTGGATATATTTCCCAAGACAGTCCAAGGT 350  
 QY 327 AGAATTTGAGGTTCTTTATGAGATTTTGAACCTGGAAGCAATGTTGCTACCTGAGATAG 386

Db 351 TGAGGTAGATGCTTATGAGAAATTGACGTTAAAAATCCTGGTAAATTACGTAGTAAG 410  
 Qy 387 CCAAGCTCAAAATTTTGGCTTGTCTCGAATGCTACAAATAGTAT- 435  
 Db 411 TGAAGCCTTCAGGTATATTGCTTTAGCAAGAGAAATTGATAATCTTCAAAAATATCCTGA 470  
 Qy 436 ----AATAAGTTATAGTACTAGAGAAATAACGGCGTTGTTGACAAAGTCTCTTAATGAAA 491  
 Db 471 AACAAATAGTATGTTGTTATAAAGAACATGCGCTTATCTGTGCAATCCATTATAATCAA 530  
 Qy 492 TGTGTTGTTAGTATGCTAGTGGTATGATTCCTTTAGCACCTTATATGCTGGTGT 551  
 Db 531 TGGCTGTATGATTTTCTTTAAACAATTTAAAGATATCACCTTTACATATGCGTAGGGTT 590  
 Qy 552 TGGTGCAGATTATATAAAGTTTTTAGGTATATCAATGCGCTTATGCTTATCAAGTTAA 611  
 Db 591 TGGTGGGCAATATATAGAAATTTTATAGTGTGCTGAAGTTTAAATTTGCTTATCAAGTTAA 650  
 Qy 612 GTTGTGTCAACTACCTCTAAATGTTAATACATATGTTGTTGGTGGGTTTATTACCA 671  
 Db 651 GGTAGTATCAGTTATCCATTATCTTAATATGATTATTTGCTGACGGATATTACCA 710  
 Qy 672 TAAGTGTAGTGTAGGATGAGAGATGAGAGATGAGAAATAGCTTACCATCTTACTGCAATTATC 731  
 Db 711 TAAGGTATAGGAAATAAATTTAAACAATTTAAATGTTCAACACGTTGTTAGTCTTAACAG 770  
 Qy 732 TGAGGTTCTAGAACTACTTACGCTTCTGCTACTTTAAATACATGATTATTTGGTGGGA 791  
 Db 771 TCA---TCCTAAGTCTACTTTGCAAGTAGCTACTCTTAATGTTGAGTATTTCCGTTAGTGA 827  
 Qy 792 GATTGGATTATGATTT 807  
 Db 828 ATTTGGTTAAATTT 843

RESULT 7  
 ABS63288  
 ID ABS63288 standard; DNA; 852 BP.  
 XX  
 AC ABS63288;  
 XX  
 DT 15-NOV-2002 (first entry)  
 XX  
 DE DNA encoding Ehrlichia chaffeensis outer membrane protein (OMP) #13.  
 XX  
 KW Outer membrane protein; OMP; gene; ds; P30F; ehrlichiosis; infection.  
 XX  
 OS Ehrlichia chaffeensis.  
 XX  
 PN US2002120115-A1.  
 XX  
 PD 29-AUG-2002.  
 XX  
 PF 28-JAN-2002; 2002US-00059964.  
 XX  
 PR 19-MAY-1999; 99US-00314701.  
 XX  
 PA (RIKI/) RIKIHISA Y.  
 PA (OHAS/) OHASHI N.  
 XX  
 PI Rikihisa Y, Ohashi N;  
 XX  
 DR WPI; 2002-618954/66.  
 DR P-PSDB; ABG77947.  
 XX  
 PT Isolated polynucleotide encoding an outer membrane protein of E.canis or  
 E.chaffeensis used in the diagnosis of infection.  
 XX  
 PS Claim 3; Fig 15A; 49pp; English.  
 XX  
 CC The invention relates to an isolated polynucleotide encoding an outer  
 CC membrane protein (OMP) of Ehrlichia canis or Ehrlichia chaffeensis used  
 CC in the diagnosis of infection. An infection such as human ehrlichiosis or

CC canine ehrlichiosis can be diagnosed by providing a serum sample from the  
 CC patient, providing a polypeptide or mixture of polypeptides, contacting  
 CC the sample with the polypeptide and assaying for the formation of a  
 CC complex between antibodies in the serum sample and the polypeptide, where  
 CC formation of a complex is indicative of infection with E. chaffeensis.  
 CC This sequence represents DNA encoding an Ehrlichia outer membrane protein  
 CC of the invention  
 XX  
 SQ Sequence 852 BP; 295 A; 115 C; 141 G; 301 T; 0 U; 0 Other;  
 Query Match 18.5%; Score 150.8; DB 6; Length 852;  
 Best Local Similarity 54.6%; Pred. No. 1e-25;  
 Matches 402; Conservative 0; Mismatches 307; Indels 27; Gaps 4;  
 Qy 96 GAGGAGAGGACTTTATATAGTAGTACAGTATATAAGTTGGTATTCCTCAATTTTAGTAATTT 155  
 Db 111 GTGCACTGGGCTATATGTGAGTGGACAATATAACCTACTGTCTCTCACTTTTAGTAATTT 170  
 Qy 156 TTCAGCTGAAGAAACAATTCCTGGTATTACAAAAAAGATTTTGGCTTAGGCTCTTGATAA 215  
 Db 171 TTCACCTTAAGAAACTTATCTGACACTTAAGAGTATTAGGACTAGCAAAAGATATTAA 230  
 Qy 216 GTCT-----GAGATAAATACCTCAGCAATTTTACACGATCATATGACCCCTACTTATGC 269  
 Db 231 GTCTATTACAGATAAACACAATAAAAAATTCACATTCCTTATAACAAAAATTTCA 290  
 Qy 270 AAGCAGTTTTCGACGGGTTTAGTGGTATCATTTGGATATTATGTTAATGACTTT--AGGGT 326  
 Db 291 AGATAATGCTGTAGCTTCAGTGCAGCTGTTGGATATATTTCCCAAGACACGTCGAAGGGT 350  
 Qy 327 AGAATTTGAAGGTTCTTATGAGAAATTTGAACCTGAAAGACAATGGTACCCTCGAGAATAG 386  
 Db 351 TGAGGTAGATGGTCTTATGAGAAATTTGACGTTAAATAATCTGTGTAATTACGTAGTAAG 410  
 Qy 387 CAAAAGCTACAAATTTTGTCTGCTAGTGGTATGTTAGCACCTTATATGTTGCTGGTGT 435  
 Db 411 TGAAGCCTTCAGGTATATGCTTTAGCAAGAGGAATTCATAATCTTCAAAAATATCCTGA 470  
 Qy 436 ----AATAAGTTTATAGTACTAGAGAAATAACGGCGTTGTTGACAAAGTCTCTTAATGTTAA 491  
 Db 471 AACAAATAGTATGTTGTTATAAAGAACAAATGGCTTATCTGTGCAATCCATTATAATCAA 530  
 Qy 492 TGTGTTGTTATGATATTGCTAGTGGTATGTTAGTATGCTTTAGCACCTTATATGTTGCTGGTGT 551  
 Db 531 TGGCTGTTATGATTTTCTTTAAACAATTTAAAGATATCACCTTACATATGCGTAGGGTT 590  
 Qy 552 TGGTGCAGATTATATAAAGTTTTTAGGTATATCATTTGCTAAGTTTTCTTTATCAAGTTAA 611  
 Db 591 TGGTGGGACATTAAGAATTTTGTAGTGTGTAAGTTTAAATTTGCTTATCAAGGTTAA 650  
 Qy 612 GTTGGTGTCAACTACCTCTAAATGTTAATGTTAATGTTGTTGGTGGGTTTATTACCA 671  
 Db 651 GGTAGGTATCAGTTATGCTTATCCATTATCTCTAATATGATTATTTGCTGACGGATATTACCA 710  
 Qy 672 TAAGGTTGTAGGTATAGGCATGAGAGATGAGAAATAGCTTACCCTACTGCAATATC 731  
 Db 711 TAAGGTATAGGAATAAATTTAAACAATTTAAATGTTCAACGTTGTTAGTCTTAACAG 770  
 Qy 732 TGACGTTCTAGAACTACTTACGCTTCTGCTACTTTTAAATACTGATTATTTGGTGGGA 791  
 Db 771 TCA---TCCTAAGTCTACTTTTGCAGTAGTACTCTTAATGTTGAGTATTTCCGTTAGTGA 827  
 Qy 792 GATTGGATTATGATTT 807  
 Db 828 ATTTGGTTAAATTT 843  
 RESULT 8  
 ADA09758  
 ID ADA09758 standard; DNA; 852 BP.  
 XX  
 AC ADA09758;  
 XX

DT	06-NOV-2003	(first entry)	
XX	E. chaffeensis	DNA encoding outer membrane protein OMP-1W.	
DE			
XX	ds; gene; outer membrane protein; circulating leukocyte;		
KW	monocytic ehrlichiosis; Rocky Mountain spotted fever;		
KW	canine ehrlichiosis.		
XX			
OS	Ehrlichia chaffeensis.		
XX			
PN	US6544517-B1.		
XX			
PD	08-APR-2003.		
XX			
PF	19-MAY-1999; 99US-00314701.		
PR	18-SEP-1998; 98US-0100843P.		
XX			
XX	(OHIS ) UNIV OHIO STATE RES FOUND.		
PA			
XX			
PI	Rikihisa Y, Ohashi N;		
XX			
DR	WPI: 2003-553952/52.		
DR	P-PSDB; ADA09759.		
XX			
PT	New isolated polynucleotide encoding outer membrane protein P30 of		
PT	Ehrlichia canis or its variant or fragment, useful for producing		
PT	Ehrlichia canis or Ehrlichia chaffeensis outer membrane proteins for		
PT	diagnosing and treating ehrlichiosis.		
XX			
PS	Disclosure; Fig 15; 105pp; English.		
XX			
CC	The invention relates to an isolated polynucleotide encoding a variant of		
CC	the outer membrane protein (OMP) P30 of Ehrlichia canis, an outer		
CC	membrane protein of E. canis , or an antigenic fragment of the E. canis		
CC	P30 protein, or comprising a sequence which is the complement of		
CC	nucleotide 99 through nucleotide 672 of ADA09764. Also disclosed are E		
CC	canis and E. chaffeensis outer membrane proteins and their encoding		
CC	nucleic acids. The polynucleotides are useful for producing E. canis or		
CC	E. chaffeensis outer membrane protein, for designing hybridisation probes		
CC	for isolating and identifying cDNA and genomic clones encoding the OMP)		
CC	or its allelic forms, for designing primers for PCR. The polypeptides		
CC	encoded by the polynucleotide is useful for diagnosing human ehrlichiosis		
CC	(Rocky mountain spotted fever) or canine ehrlichiosis. The present		
XX	sequence encodes an E. chaffeensis outer membrane protein.		
XX			
SQ	Sequence 852 BP; 295 A; 115 C; 141 G; 301 T; 0 U; 0 Other;		
Query Match 18.5%; Score 150.8; DB 9; Length 852;			
Best Local Similarity 54.6%; Pred. No. 1e-25;			
Matches 402; Conservative 0; Mismatches 307; Indels 27; Gaps 4			
Qy	96	GAGGAGAGGACHTTATATAGTAGTCAGTATAAGTTGGTATTTCCCAATTTTAGTAATTT 155	
Db	111	GTGCATCGGGCTATATGTGTCAGTGGACATATAAACCCTACTGTTCTCAGCTTAGTAATTT 170	
Qy	156	TTCCAGCTGAAGAAACAAATTCCTGGTATTACAAAAAGATTTTTCGGTTAGGTCCTTGATAA 215	
Db	171	TTCCCTTAAAGAAACTTATCTATGACATTAAGAGTTATTAGGACTAGCAAAAGATATTAA 230	
Qy	216	GTCT-----GAGATAAAATCTCAGACAAATTTTACACGATCATATGACCCCTACTTATGC 269	
Db	231	GTCTATTACAGATATAACAAACAATAAAAAATTCACATTCCTTATAACACAAAATTTCA 290	
Qy	270	AGCAGTTTTGCAGGGTTTAGTGGTATCAATTTGGATATATATGTAAGACATTT---AGGCT 326	
Db	291	AGATAATGCTGTTAGCTTCAGTGCAGCTGTTGGATATATTTCCCAAGACAGTCCAAGGGT 350	
Qy	327	AGAAATTTGAAGTCTTATGAGAAATTTTGAACCTGGAAGACAATGGTACCCTCGAGAATAG 386	
Db	351	TGAGGTAGAAATGGTCTTATGAGAAATTTGACGTTAAAAAATCCCTGGTAATTTACGTAGTAAG 410	
Qy	387	CCAAAGCTACAAATTTTTTGTCTTCGAAATGCTACAAATAGTGAT----- 435	



```
Db 301 ATATCAGGATTTTCAGGAAGTATGGTTACTCTATGACGACGACCAAGATAGAACTTGAA 360
Qy 337 GGTTCCTTATGAGAAATTTTGAACCTGGAAGACAATGGTACC-----CTGAGAATAGCCAA 390
Db 361 GCTGCATATCAACAATTTAATCCAAAAACACCGGATAACAATGATATCTGATAATGGTGAA 420
Qy 391 AGCTACAAAATTTTGTCTGCGAAATGCTACAAATAGTGAATAATAGTTTATAGTA 450
Db 421 TACTATAAACATTTTGCATTATCTCGTAAAGATGCAATGGAAGATCAGCAATATGTAGTA 480
Qy 451 CTACAGAAATACGCGGTGTTGACAACTCTTAATGTAATGTTTCTGTTATGATATTGCT 510
Db 481 CTTAAATAGCGGCAATCTTTATGTCATTGATGGTTAATCTCTCTATGACATTACA 540
Qy 511 AGTGGTAGTATTCCTTTAGCACCTTATATATGCTGCTGCTGTTGGTGAGATTTATATAAG 570
Db 541 GCTGAAGGAGTATCTTTTCGTACCATATGCGATGTCAGGTATAGGAGCAGATCTTATCACT 600
Qy 571 TTTTGGGTATATCAATTCGCTTAAGTCTTATCAAGTTTCTTATCAAGTTTGGTGCACTACCT 630
Db 601 ATTTTAAAGACCTCAATCTAAAAATTTGCTTTACCAAGGAAAAATAGGTATTTAGTTACCT 660
Qy 631 CTAAATGTTAATACTATGTTGTTGGTGGGTTTATTACCATAGGTTTGGTAGGTATAGG 690
Db 661 ATCACACAGAGTCTCTCGCAATTTATTGGTGGGATCTACCATGCGGTTATTTGGTAATAA 720
Qy 691 CATGAGAGAGTAGAAATAGCTTACCATCTCTATGCAATTTCTGACGTTCTCTAGAACTACT 750
Db 721 TTTGAGAAGATACCTGTA---ATAACTCTCTGTAGTATTTAAATGATGCTCTCTCAACCCACA 777
Qy 751 TCAGCTTCTGCTACTTTTAAATACATGATTTATTTGGTTGGGAGATTGGATTAGATT 806
Db 778 TCTGCTTCAGTAACTCTTTGACGTTGGATCTTTGGCGGAGAAATTTGGAATGAGGTT 833
```

## RESULT 11

```
ABS63299
ID ABS63299 standard; DNA; 843 BP.
AC ABS63299;
XX
DT 15-NOV-2002 (first entry)
XX
DE DNA encoding Ehrlichia canis outer membrane protein (P30F) #9.
XX
KW Outer membrane protein; OMP; gene; ds; P30F; ehrlichiosis; infection.
XX
OS Ehrlichia canis.
XX
PN US2002120115-A1.
XX
PD 29-AUG-2002.
XX
PF 28-JAN-2002; 2002US-00059964.
XX
PR 19-MAY-1999; 99US-00314701.
XX
PA (RIKI/) RIKIHISA Y.
PA (OHAS/) OHASHI N.
XX
PI Rikihisa Y, Ohashi N;
XX
DR WPI; 2002-618954/66.
DR P-PSDB; ABG77958.
XX
```

```
XX Isolated polynucleotide encoding an outer membrane protein of E.canis or
PT E.chaffeensis used in the diagnosis of infection.
XX
PS Claim 5; Fig 30A; 49pp; English.
XX
CC The invention relates to an isolated polynucleotide encoding an outer
CC membrane protein (OMP) of Ehrlichia canis or Ehrlichia chaffeensis used
```

```
CC in the diagnosis of infection. An infection such as human ehrlichiosis or
CC canine ehrlichiosis can be diagnosed by providing a serum sample from the
CC patient, providing a polypeptide or mixture of polypeptides, contacting
CC the sample with the polypeptide and assaying for the formation of a
CC complex between antibodies in the serum sample and the polypeptide, where
CC formation of a complex is indicative of infection with E. chaffeensis.
CC This sequence represents DNA encoding an Ehrlichia outer membrane protein
CC of the invention
XX
```

SQ Sequence 843 BP; 303 A; 141 C; 144 G; 255 T; 0 U; 0 Other;

```
Query Match 18.1%; Score 147.2; DB 6; Length 843;
Best Local Similarity 53.2%; Pred. No. 7.4e-25;
Matches 445; Conservative 0; Mismatches 358; Indels 33; Gaps 5;

Qy 1 ATGAATTAACAAAGATTTGTGTAGGTGTACGCTGAGTACATTTGTTTTTCTTATCT 60
Db 1 ATGAATTAACAAAGATTTGTGTAGGTGTACGCTGAGTACATTTGTTTTTCTTATCT 60
Qy 61 GATGGTCTTTTCTGA-----TGCATAATTTTCTGAAGGAGGAGGACTTTAT 111
Db 61 TATCAGTCTTTTTCGAGATCCTGTAGGTTCAGAACTAATGATAACAAAGAGGCTTCTAC 120
Qy 112 ATAGGTAGTCAGTATAAAGTTGGTATTTCCCAATTTTAGTAATTTTTCAGCTGGAAGAAACA 171
Db 121 ATTAGTGCAAGTACAACTCCAGTATATCACCTTTAGAAAAATTCCTGCTGGAAGAACT 180
Qy 172 ATTCTCTGGTATTAACAAAAAGATTTTTCGCTTGGCTTGGTCTTGG-----ATAAGTCTGAG 222
Db 181 CCTATTAAATGGAACAAATCTCTCACTAAAAAGTTTTCGGACTAAAGAAAGATGGTGAT 240
Qy 223 ATAAATACTCACAGCAATTTTACAGTACATATGACCCCTACTTAT-----GCAAGCAGT 276
Db 241 ATAACAAAAAAGACGATTTTACAGAGTAGTCTCCAGGCAATGATTTTCAAAAATAACTTA 300
Qy 277 TTTGCGAGGTTTGTAGTGTATCATTTGATATTTTGTAAATGATGATTTAGGTTAGAAATTTGAA 336
Db 301 ATATCAGGATTTTCAGGAAGTATTTGTTACTCTATGGACGACCAAGATAGAACTTGAA 360
Qy 337 GGTTCCTTATGAGAAATTTTGAACCTGGAAGACAATGGTACC-----CTGAGAATAGCCAA 390
Db 361 GCTGCATATCAACAATTTTAAATCCAAAAACACCGATAACAATGATGATGATGATGATGAA 420
Qy 391 AGCTACAAATTTTGTCTTCTCGAAATGCTACAAATAGTGTATGATATAGTATAGTA 450
Db 421 TACTATAAACAATTTTGCATTATCTCGTAAAGATGCAATGGAAGATCAGCAATATGTAGTA 480
Qy 451 CTAGAGAAATAACGCGGTTGTTGACAAGTCTCTTAATGTAATGTTTCTGTTATGATATTGCT 510
Db 481 CTTAAATAGTACGCGCATAACTTTTATGTCATTGATGGTTAATCTGCTCTATGACATTACA 540
Qy 511 AGTGGTAGTATTCCTTTAGCACCTTATATGCTGCTGCTGTTGGTGAGATTTATATAAG 570
Db 541 GCTGAAGGAGTATCTTTTCGTACCATATGCGATGTCAGGTATAGGAGCAGATCTTATCACT 600
Qy 571 TTTTGGGTATATCAATTCGCTTAAGTCTTCTTATCAAGTTTGGTTGGTCTCACTACCT 630
Db 601 ATTTTAAAGACCTCAATCTAAAAATTTGCTTTACCAAGGAAAAATAGGTATTTAGTTACCT 660
Qy 631 CTAAATGTTAATACTATGTTTGGTGGGTTTATTACCATAGGTTTGGTAGGTATAGG 690
Db 661 ATCACACAGAGTCTCTCGCAATTTATTGGTGGGATCTACCATGCGGTTATTTGGTAATAA 720
Qy 691 CATGAGAGAGTAGAAATAGCTTACCATCTCTATGCAATTTCTGACGTTCTCTAGAACTACT 750
Db 721 TTTGAGAAGATACCTGTA---ATAACTCTCTGTAGTATTTAAATGATGCTCTCTCAACCCACA 777
Qy 751 TCAGCTTCTGCTACTTTTAAATACATGATTTATTTGGTTGGGAGATTGGATTAGATT 806
Db 778 TCTGCTTCAGTAACTCTTTGACGTTGGATCTTTGGCGGAGAAATTTGGAATGAGGTT 833
```

RESULT 12

ADA09780  
ID ADA09780 standard; DNA; 843 BP.  
AC ADA09780;  
XX  
XX 06-NOV-2003 (first entry)  
DT  
XX  
XX E. canis DNA encoding outer membrane protein p30-10.  
DE  
XX ds; gene; outer membrane protein; circulating leukocyte;  
KW monocytic ehrlichiosis; Rocky Mountain spotted fever;  
KW canine ehrlichiosis.  
XX  
XX Ehrlichia canis.  
XX  
XX US6544517-B1.  
PN  
XX  
XX  
XX 08-APR-2003.  
PD  
XX  
XX 19-MAY-1999; 99US-00314701.  
PF  
XX  
XX 18-SEP-1998; 98US-0100843P.  
PR  
XX  
XX (OHIS ) UNIV OHIO STATE RES FOUND.  
PA  
XX  
XX Rikihisa Y, Ohashi N;  
PI  
XX  
XX WPI; 2003-553952/52.  
DR  
XX  
XX P-PSDB; ADA09781.  
DR  
XX  
XX  
XX New isolated polynucleotide encoding outer membrane protein P30 of  
PT Ehrlichia canis or its variant or fragment, useful for producing  
PT Ehrlichia canis or Ehrlichia chaffeensis outer membrane proteins for  
PT diagnosing and treating ehrlichiosis.  
XX  
XX  
XX Disclosure; Fig 30; 105pp; English.  
PS  
XX  
XX The invention relates to an isolated polynucleotide encoding a variant of  
CC the outer membrane protein (OMP) P30 of Ehrlichia canis, an outer  
CC membrane protein of E. canis , or an antigenic fragment of the E. canis  
CC P30 protein, or comprising a sequence which is the complement of  
CC nucleotide 99 through nucleotide 672 of ADA09764. Also disclosed are E.  
CC canis and E. chaffeensis outer membrane proteins and their encoding  
CC nucleic acids. The polynucleotides are useful for producing E. canis or  
CC E. chaffeensis outer membrane protein, for designing hybridisation probes  
CC for isolating and identifying cDNA and genomic clones encoding the OMP)  
CC or its allelic forms, for designing primers for PCR. The polypeptides  
CC encoded by the polynucleotide is useful for diagnosing human ehrlichiosis  
CC (Rocky mountain spotted fever) or canine ehrlichiosis. The present  
CC sequence encodes an E. canis outer membrane protein.  
XX  
XX  
XX Sequence 843 BP; 303 A; 141 C; 144 G; 255 T; 0 U; 0 Other;  
SQ  
  
Query Match 18.1%; Score 147.2; DB 9; Length 843;  
Best Local Similarity 53.2%; Pred. No. 7.4e-25;  
Matches 445; Conservative 0; Mismatches 358; Indels 33; Gaps 5;  
  
QY 1 ATCAATATCAAAAGATTGTTAGGTGTACCTCAGTACATTTGTTTCTTATCT 60  
Db 1 ATCAATATCAAAAGATTGTTAGGTGTACCTCAGTACATTTGTTTCTTATCT 60  
QY 61 GATGGTGTCTTTTCTGA-----TGCAAAATTTTCTGAAGGGAGGAGACTTTAT 111  
Db 61 TATCAGTCTTTTGCAGATCTGTAGGTTCAGAACTAATGATAACAAGAGGCTTCTAC 120  
QY 112 ATAGGTAGTCAGTATAAAGTTGGTATGCCAATTTTAGTAAATTTTTCAGCTGAAGAAACA 171  
Db 121 ATTAGTGCAAGGTACAACTCCAGTATATCACACTTTAGAAAATTTCTCTGCTGAAGAACT 180  
QY 172 ATTCCTGGTATTACAAAAGATTTTTCGTTAGGTCTTG-----ATAAGTCTGAG 222  
Db 181 CCTATTAAATGGAACAAATTTCTCTCACTAAAAAAGTTTTCGGACTAAAGAAAGATGGTAT 240

223 ATAAATACTCACAGCAATTTTACAGATCATATGACCCCTACTTAT-----GCAAGCAGT 276  
Db 241 ATAACAAAAAAGAGAGATTTTACAAGAGTAGCTCCAGGCATTGATTTTCAAAATAACTTA 300  
QY 277 TTTCAGGGTTTAGTGGTATCATTTGGATATTTATGTTAATGACTTTAGGGTAGAATTTGAA 336  
Db 301 ATATCAGGATTTTTCAGGAAAGTATTTGGTTACTCTATGGACCGACCAAGAAATAGAATTTGAA 360  
QY 337 GGTCTTATGAGAAATTTTGAACCTGAAAGACAAATGGTACC-----CTGAGAATAGCCAA 390  
Db 361 GCTGCATATCAAAATTTTAAATCCAAAAAACACCGATAACAATGATGACTGATAATGCTGAA 420  
QY 391 AGCTACAAATTTTGTCTCGAAATGCTCAAAATAGTGATAATAAGTTTATAGTA 450  
Db 421 TACTATAAACAATTTTGCATTTATCTCGTAAAGATGCAATGGAAGATCAGCAATATCTAGTA 480  
QY 451 CTAGAGAATAACGGCGTTTGTGACAAAGTCTCTTAATGTAATGTTTGTGTTATGATATTCGT 510  
Db 481 CTTAAAAATGACGGCATAACTTTTATGTCATTGATGGTTAATGCTATGCTATGACATTACA 540  
QY 511 AGTGGTAGTATTCCTTTAGCACCTTATATGTCGTGCTGTTGGTGCAGATTATATAAG 570  
Db 541 GCTGAAGGAGTATCTTTCTGACCATATGATGTCAGGTATAGGAGCAGATCTTATCACT 600  
QY 571 TTTTATAGGTATATCATTCCTTAAGTTTCTTATCAAGTTAAAGTTTGGTGTCAACTACCT 630  
Db 601 ATTTTAAAGACCTCAATCTAAATTTTCTTACCAAGGAAAAATAGGTATTAGTTACCT 660  
QY 631 CTAAATGTTAATACTATGTTGTTGGTGGGGTTTATACCATAGAGTTGTAGGTGATAGG 690  
Db 661 ATCACCAGAGTCTCTGTCATTTATTTGGTGGATACCTACCATGGCGTTTATTTGGTAAATAA 720  
QY 691 CATGAGAGAGTAGAAATAGCTTACCATCTACTGCTATCTGACGTTCTCTAGAACTACT 750  
Db 721 TTTGAGAAGATACCTGTA---ATAACTCTGTAGTATTAAATGATGCTCTCTCAAAACCACA 777  
QY 751 TCAGCTTCTGCTACTTTAAATCTGATTATTTTGGTGGGAGATTGGATTAGATT 806  
Db 778 TCTGCTTCAGTAACTCTTGACGTTGGATACCTTTGGCGGAGAAATTTGGAATGAGTT 833  
  
RESULT 13  
ADW04273  
ID ADW04273 standard; DNA; 843 BP.  
XX  
XX AC ADW04273;  
XX  
XX 24-MAR-2005 (first entry)  
DT  
XX  
XX Ehrlichia canis outer membrane protein (P30F) DNA, P30-10.  
XX  
XX DNA purification; diagnosis; outer membrane protein; OMP; P30F protein;  
KW infection; vaccine; gene; ds.  
XX  
XX Ehrlichia canis.  
OS  
XX  
XX Key Location/Qualifiers  
FT CDS 1..843  
FT /tag= b  
FT /product= "Ehrlichia canis outer membrane protein (OMP)"  
FT sig\_peptide 1..75  
FT /tag= a  
FT mat\_peptide 27..840  
FT /tag= c  
FT /product= "Ehrlichia canis mature OMP protein"  
XX US2004265334-A1.  
XX  
XX 30-DEC-2004.  
PD  
XX 29-JUL-2004; 2004US-00901774.  
PF  
XX 18-SEP-1998; 98US-0100843P.  
PR



PR 19-MAY-1999; 99US-00314701.  
 PR 28-JAN-2002; 2002US-00059964.  
 XX (RIKI/) RIKIHISA Y.  
 PA (OHAS/) OHASHI N.  
 XX Rikihisa Y, Ohashi N;  
 PI WPI; 2005-064871/07.  
 DR P-PSDB; ADM04274.  
 XX New polynucleotide encoding an outer membrane protein (OMP) of Ehrlichia  
 PT canis or Ehrlichia chaffeensis, useful in preparing a composition for  
 PT diagnosing or preventing E. canis or E. chaffeensis infection.  
 XX  
 PS Claim 2; SEQ ID NO 47; 122pp; English.  
 XX  
 CC The invention relates to nucleic acid sequences encoding outer membrane  
 CC proteins (OMP) of Ehrlichia chaffeensis (designated as OMP proteins) and  
 CC Ehrlichia canis (designated as P30F proteins). The OMP polynucleotide is  
 CC useful in preparing a composition for diagnosing, treating or preventing  
 CC an infection with Ehrlichia canis or Ehrlichia chaffeensis. The present  
 CC sequence is the Ehrlichia canis P30F DNA.  
 XX  
 XX Sequence 843 BP; 303 A; 141 C; 144 G; 255 T; 0 U; 0 Other;  
 SQ  
 Query Match 18.1%; Score 147.2; DB 14; Length 843;  
 Best Local Similarity 53.2%; Pred. No. 7.4e-25;  
 Matches 445; Conservative 0; Mismatches 358; Indels 33; Gaps 5;  
 QY 1 ATGAATACAAAAGATTGTTGGTAGGTGTACCGCTGAGTACATTTGTTTCTTATCT 60  
 DB 1 ATGAATATAGAAGAAATCTAGTAGAAGCGGTAAATCTCATTTAATGTCAATCTTACCA 60  
 QY 61 GATGGTCTTTTCTGA-----TGCRAATTTTCTGAAGGAGGAGGACTTTAT 111  
 DB 61 TATCAGTCTTTTGCAGATCCTGTAGGTTCAAGAATTAANGATACAAGAGGCTTCTAC 120  
 QY 112 ATAGGTAGTCAGTAAAGTTGGTATTCCTCAATTTAGTAAATTTTTCAGCTGAAGAAACA 171  
 DB 121 ATTAGTGCAGATCAATCCAGATATATACACATTTAGAAATTTCTCTGCTGAAGAACT 180  
 QY 172 ATTCCTGGTATACAAAAGATTTTTCGGTGTAGGTCTG-----ATAAGTCTGAG 222  
 DB 181 CCTATTAAATGCAAAATTTCTCTCACTAAAGAAAGTTTCGGACTAAAGAAAGATGGTAT 240  
 QY 223 ATAAATCTCAGCAATTTTACAGATCATATGACCTCTCTAT-----GCAAGCAGT 276  
 DB 241 ATAAACAAAAGACGATTTTACAGAGTAGCTCCAGGCATTTGATTTTCAAAATAACTTA 300  
 QY 277 TTTGAGGTTTGTAGGTATCATTTGGATATTTATGATGATTTTGTAGGTAGAAATTTGAA 336  
 DB 301 ATATCAGGATTTTCAGGAAGTATTTGTTACTCTATGGAGGACCAAGATAGAACTTGA 360  
 QY 337 GGTCTTATAGAAATTTTGAACCTGGAAGACAAATGGTACC-----CTGAGAAATAGCCAA 390  
 DB 361 GCTGCATATCAACATTTTAAATCCAAAGAACCCGATACATGATGATAAATGGTGA 420  
 QY 391 AGCTACAAATTTTGTCTTCTCGAATGCTCAATATGATGATGATGATGATGATGATGAT 450  
 DB 421 TACTATAACAAATTTTGCATATCTGTAAAGATGCAATGGAAGATCAGCAATATGTAGTA 480  
 QY 451 CTAGAGATAACCGGTTTGTGACAAAGTCTCTTAATGTAAATTTTGTATGATATTTGCT 510  
 DB 481 CTTAAANATGCGGCATAACTTTATGTCATTTGATGTTTATGTTTATGTTTATGATGAT 540  
 QY 511 AGTGTAGTATTCCTTTAGACCTTATATGTTGCTGTGTTGTTGTTGTTGTTGTTGTTGTT 570  
 DB 541 GCTGAAGGAGTATCTTTTCTGATCATATGCTGAGGTATGAGGAGGAGATCTTTATCACT 600  
 QY 571 TTTTATAGGTATATCATTCCTTAAGTTTCTTATCAAGTTTAAAGTTTGGTCACTACCT 630  
 DB 601 ATTTTAAAGACCTCAATCTAAATTTTGCCTTACCAGGAAATATAGGTATTTAGTTTACCCT 660

QY 631 CTAATGTTAATACTATGTTTGGTGGGGTTATTACCATAGCTTGTAGGTATAGG 690  
 DB 661 ATCACACCAAGAGTCTCTGCAATTTATTTGGTGATACCTACCATGGCGTTATTTGTAATAAA 720  
 QY 691 CATGAGAGAGTAGAATAAGCTTACCATCTCTACTGCAATTTATCTGACGTTCTCTAGAACTACT 750  
 DB 721 TTTGAGAGATACCTGTA--ATACTCTCTGTAGTATTAATGATGCTCTCTCAACACCA 777  
 QY 751 TCAGCTTCTGCTACTTTAAATACTGATTTATTTGGTTGGGAGATTTGATTTAGTT 806  
 DB 778 TCTGCTTCACTAATCTCTGACGTTGGATATTTTGGCGGAGAAATTTGGAATGAGTT 833  
 RESULT 14  
 ABS63306  
 ID ABS63306 standard; DNA; 726 BP.  
 XX  
 AC ABS63306;  
 XX  
 DT 15-NOV-2002 (first entry)  
 XX  
 DE DNA encoding Ehrlichia canis outer membrane protein (P30F) #14.  
 XX  
 KW Outer membrane protein; OMP; gene; ds; P30F; ehrlichiosis; infection.  
 XX  
 OS Ehrlichia canis.  
 XX  
 PN US2002120115-A1.  
 XX  
 PD 29-AUG-2002.  
 XX  
 PF 28-JAN-2002; 2002US-00059964.  
 XX  
 PR 19-MAY-1999; 99US-00314701.  
 XX  
 PA (RIKI/) RIKIHISA Y.  
 PA (OHAS/) OHASHI N.  
 XX  
 PI Rikihisa Y, Ohashi N;  
 XX  
 DR WPI; 2002-618954/66.  
 DR P-PSDB; ABG77965.  
 XX  
 PT Isolated polynucleotide encoding an outer membrane protein of E. canis or  
 PT E. chaffeensis used in the diagnosis of infection.  
 XX  
 PS Disclosure; Fig 32A; 49pp; English.  
 XX  
 CC The invention relates to an isolated polynucleotide encoding an outer  
 CC membrane protein (OMP) of Ehrlichia canis or Ehrlichia chaffeensis used  
 CC in the diagnosis of infection. An infection such as human ehrlichiosis or  
 CC canine ehrlichiosis can be diagnosed by providing a serum sample from the  
 CC patient, providing a polypeptide or mixture of polypeptides, contacting  
 CC the sample with the polypeptide and assaying for the formation of a  
 CC complex between antibodies in the serum sample and the polypeptide, where  
 CC formation of a complex is indicative of infection with E. chaffeensis.  
 CC This sequence represents DNA encoding an Ehrlichia outer membrane protein  
 CC of the invention  
 XX  
 SQ Sequence 726 BP; 255 A; 99 C; 112 G; 260 T; 0 U; 0 Other;  
 Query Match 18.0%; Score 146.6; DB 6; Length 726;  
 Best Local Similarity 54.0%; Pred. No. 9.8e-25;  
 Matches 358; Conservative 0; Mismatches 284; Indels 21; Gaps 2;  
 QY 166 GAACAATTCCTGGTATTACAAAAGATTTTTCGGTTAGGTCTTGATAGTCTGAGATA 225  
 DB 55 GAGGCATTTGTTGGCTTAAACCAAGATATTAGTTCTATTATTTTACGTAATAAGAGACCA 114  
 QY 226 AATACTCACAGCAATTTTACACGATCATATGACCTACTTATGCAAGCAGTTTTCAGAGG 285  
 DB 115 CAATATTAACAATTTTAAAGTTCCCTCTATCTGCAAAATTTCAAGAGCAGCTTTGCGAGT 174

QY 286 TTTAGTGGTATCATTTGGATATTATGTT---AATGACTTTAGGGTAGAAATTTGAAGTTCT 342  
DB 175 TTCCAGCATAGCTGTTGGATATATTGCTTAACAATGGTCCAGAAATTGAAATAGAGGATCT 234  
QY 343 TATGAGAAATTTGAACCTGAAAGACAATGGTACCCTGGAGATAGCCAAAGCTACAAATTT 402  
DB 235 TACGAAGAAATTTGATGTTAAACCCAGGAAATTATACAACAATAGATGCTCATGGTAC 294  
QY 403 TTTCCTTTGCTCGAAATGCTACAAA-----TAGTGATTAATAGTTT 444  
DB 295 ATTGCTTTAGCTAGAGAAAAAATCTTCTTACTATCTAAGTTCTCTTAAAGAAAAACAAATAT 354  
QY 445 ATAGTACTAGAGAAATACCGCGTTGTTGACAACTCTTAATGCTAAATGTTTGTATGAT 504  
DB 355 GTAAATTAAGAAATTAACGGCATATCTATTGTAICTATTATTAATTAATGTTGTTAATGAT 414  
QY 505 ATTGCTAGTGGTAGTATTCTTTTAGCACCTTATATGCTGCTGTTGTTGGTGAGATTAT 564  
DB 415 ATTTCTTTAAATGATTCTTAAGGTGTCACTTACATATGCACAGGGTTTGGTGGAGATTT 474  
QY 565 ATAAAGTTTATAGTATATCATTTGCTTAAGTTTCTTATCAAGTTAAAGTTTGGTGTCAAC 624  
DB 475 ATAGAGTTTATAGTGTATACGTTTAAAGTTTGGTTTATCAAGGTAATAATAGTATCAGT 534  
QY 625 TACCTCTAAATGTTTAATACTATGTTTGGTGGGGTTATTACCAATAGGTTGTAGCT 684  
DB 625 TACCTCTAAATGTTTAATACTATGTTTGGTGGGGTTATTACCAATAGGTTGTAGCT 684  
QY 535 TATTCAATTATCTTCAACATAATTTTATTACTGATGGATATTAACCAAGGTAATAAAT 594  
QY 685 GATAGGATGAGAGAGTAGAAATAGCTTACCATCTCTACTGCAATATCTGACGTTCCCTAGA 744  
DB 595 TCCCAATTTAAAGAAATTTAAATGTTGAACATGTTGTTAAATGAGTTAACTACAGATCCTAAA 654  
QY 745 ACTACTTCAGCTTCTGCTACTTTAAATATCTGATTTATTTGGTGGAGATTTCGATTTAGA 804  
DB 655 GTGACTTCTGCAACAGCATTTCTTAATATTGAGTATTTTGGTGGTGAATTTGGATTAAAA 714  
QY 805 TTT 807  
DB 715 TTT 717

RESULT 15  
ADA09794  
ID ADA09794 standard; DNA; 726 BP.  
XX  
AC ADA09794;  
XX  
XX 06-NOV-2003 (first entry)  
XX  
XX E. canis DNA encoding outer membrane protein p30-12.  
XX  
XX db; gene; outer membrane protein; circulating leukocyte;  
XX monocyctic ehrlichiosis; Rocky Mountain spotted fever;  
XX canine ehrlichiosis.  
XX  
XX Ehrlichia canis.  
XX  
XX US6544517-B1.  
XX  
XX 08-APR-2003.  
XX  
XX 19-MAY-1999; 99US-00314701.  
XX  
XX 18-SEP-1998; 98US-0100843P.  
XX  
XX (OHIS ) UNIV OHIO STATE RES FOUND.  
XX  
XX Rikihisa Y, Ohashi N;  
XX  
XX WPI; 2003-553952/52.  
XX  
XX P-PSDB; ADA09795.  
XX

PT New isolated polynucleotide encoding outer membrane protein p30 of  
PT Ehrlichia canis or its variant or fragment, useful for producing  
PT Ehrlichia canis or Ehrlichia chaffeensis outer membrane proteins for  
XX diagnosing and treating ehrlichiosis.  
XX  
PS Disclosure; Fig 32; 105pp; English.  
XX  
CC The invention relates to an isolated polynucleotide encoding a variant of  
CC the outer membrane protein (OMP) P30 of Ehrlichia canis, an outer  
CC membrane protein of E. canis , or an antigenic fragment of the E. canis  
CC P30 protein, or comprising a sequence which is the complement of  
CC nucleotide 99 through nucleotide 672 of ADA09764. Also disclosed are E.  
CC canis and E. chaffeensis outer membrane proteins and their encoding  
CC nucleic acids. The polynucleotides are useful for producing E. canis or  
CC E. chaffeensis outer membrane protein, for designing hybridisation probes  
CC for isolating and identifying cDNA and genomic clones encoding the OMP  
CC or its allelic forms, for designing primers for PCR. The polypeptides  
CC encoded by the polynucleotide is useful for diagnosing human ehrlichiosis  
CC (Rocky mountain spotted fever) or canine ehrlichiosis. The present  
CC sequence encodes an E. canis outer membrane protein.  
XX  
SQ Sequence 726 BP; 255 A; 99 C; 112 G; 260 T; 0 U; 0 Other;  
Query Match 18.0%; Score 146.6; DB 9; Length 726;  
Best Local Similarity 54.0%; Pred. No. 9.8e-25;  
Matches 358; Conservative 0; Mismatches 284; Indels 21; Gaps 2;  
QY 166 GAAACAATTCCTGGTATTACAAAAAGATTTTTTGGGTTAGTCTTGTGATAAGTCTGAGATA 225  
DB 55 GAGGCATTTGTTGGGCTAAAAACAGATATTAGTTCTATTATTCGTAATAAGAGACCACA 114  
QY 226 AATCTCTACAGCAATTTTACAGCATCATATGACCTCTATGCAAGCAGTTTTCGACGG 285  
DB 115 CAATATAATAACAATTTTAAAGTTCCCTATCTGCAAAATTTCAAGACGACTTTGGAGT 174  
QY 286 TTTAGTGTATCATTTGGATATTATGTT---AATGACTTTTAGGGTAGAATTTGAAGGTTCT 342  
DB 175 TTCAGCATAGCTGTTGGATATATTGCTAAACAATGGTCCAGAAATTTGAAATAGAGATCT 234  
QY 343 TATGAGAAATTTGAACCTGAAAGACAATGGTACCCTGAGATAGCCAAAGCTACAAATTT 402  
DB 235 TACGAAGAAATTTGATGTTAAAGAAACCCAGGAAATTTATACAACAATAGATGCTCATAGTAC 294  
QY 403 TTTCCTTTGCTCGAAATGCTACAAA-----TAGTGATAATAAGTTT 444  
DB 295 ATTGCTTTAGCTAGAGAAAAAATCTTCTTACTATCTAAGTTCTCTTAAAGAAAAACAAATAT 354  
QY 445 ATAGTACTAGAGAAATACCGCGTTGTTGACAAAGTCTCTTAATGTAATGTTTGTATGAT 504  
DB 355 GTAAATTAAGAAATTAACGGCATATCTATTGTAICTATTATTAATTAATGTTTGTATGAT 414  
QY 505 ATTGCTAGTGGTAGTATTCTTTTAGCACCTTATATGCTGCTGTTGTTGGTGGCAGATTAT 564  
DB 415 ATTTCTTTAAATGATTCTTAAGGTGTCACTTACATATGCACAGGGTTTGGTGGAGATTTT 474  
QY 565 ATAAAGTTTATAGGATATATCATTTGCTTAAGTTTCTTATCAAGTTAAAGTTTGGTGTCAAC 624  
DB 475 ATAGAGTTTATAGTGTATACGTTTAAAGTTTGGTTTATCAAGGTAATAATAGTATCAGT 534  
QY 625 TACCTCTAAATGTTTAATACTATGTTTGGTGGGGTTATTACCAATAGGTTGTAGCT 684  
DB 535 TATTCAATTATCTTCAACATAATTTTATTACTGATGGATATTACCAAGGTAATAAAT 594  
QY 685 GATAGGATGAGAGAGTAGAAATAGCTTACCATCTCTACTGCAATATCTGACGTTCCCTAGA 744  
DB 595 TCCCAATTTAAAGAAATTTAAATGTTGAACATGTTGTTAAATGAGTTAACTACAGATCCTAAA 654  
QY 745 ACTACTTCAGCTTCTGCTACTTTAAATATCTGATTTATTTGGTGGAGATTTCGATTTAGA 804  
DB 655 GTGACTTCTGCAACAGCATTTCTTAATATTGAGTATTTTGGTGGTGAATTTGGATTAAAA 714  
QY 805 TTT 807  
DB 715 TTT 717

PT New isolated polynucleotide encoding outer membrane protein p30 of  
PT Ehrlichia canis or its variant or fragment, useful for producing  
PT Ehrlichia canis or Ehrlichia chaffeensis outer membrane proteins for  
XX diagnosing and treating ehrlichiosis.  
XX  
PS Disclosure; Fig 32; 105pp; English.  
XX  
CC The invention relates to an isolated polynucleotide encoding a variant of  
CC the outer membrane protein (OMP) P30 of Ehrlichia canis, an outer  
CC membrane protein of E. canis , or an antigenic fragment of the E. canis  
CC P30 protein, or comprising a sequence which is the complement of  
CC nucleotide 99 through nucleotide 672 of ADA09764. Also disclosed are E.  
CC canis and E. chaffeensis outer membrane proteins and their encoding  
CC nucleic acids. The polynucleotides are useful for producing E. canis or  
CC E. chaffeensis outer membrane protein, for designing hybridisation probes  
CC for isolating and identifying cDNA and genomic clones encoding the OMP  
CC or its allelic forms, for designing primers for PCR. The polypeptides  
CC encoded by the polynucleotide is useful for diagnosing human ehrlichiosis  
CC (Rocky mountain spotted fever) or canine ehrlichiosis. The present  
CC sequence encodes an E. canis outer membrane protein.  
XX

Db 715 TTT 717

Search completed: December 28, 2005, 10:56:44  
Job time : 626 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 28, 2005, 10:46:12 ; Search time 4129 Seconds  
(without alignments)  
9212.375 Million cell updates/sec

Title: US-10-731-554-45  
Perfect score: 813  
Sequence: 1 atgaatacaaaagattgt.....ttggatttagattggcta 813

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues  
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_est3.\*  
4: gb\_est4.\*  
5: gb\_est5.\*  
6: gb\_est6.\*  
7: gb\_est7.\*  
8: gb\_est8.\*  
9: gb\_gsa1.\*  
10: gb\_gsa2.\*  
11: gb\_gsa3.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	62.8	7.7	1101	10	CNS0039G	AL063921 Drosophil
C 2	60.8	7.5	1101	10	CNS0100X	AL098379 Drosophil
C 3	56.8	7.0	829	10	EX173672	EX173672 Danio rer
C 4	56.4	6.9	1101	10	CNS00EVL	AL069706 Drosophil
C 5	55.8	6.9	1101	10	CNS00ESI	AL069797 Drosophil
C 6	54.2	6.7	1101	10	CNS00LOO	AL068607 Drosophil
C 7	54	6.6	530	7	CK573092	CK573092 1099 Plas
C 8	53	6.5	1101	10	CNS00LT2	AL078714 Drosophil
C 9	52.6	6.5	976	10	CNS010DV	AL098845 Drosophil
C 10	51.2	6.3	630	10	EX152436	EX152436 Danio rer
C 11	51	6.3	844	10	EX139987	EX139987 Danio rer
C 12	51	6.3	889	10	CW973509	CW973509 AIAA-aaa5
C 13	50.4	6.2	1094	10	CNS012PZ	AL101513 Drosophil
C 14	50.2	6.2	471	10	CG807194	CG807194 1118078H0
C 15	50	6.2	670	9	BZ388480	BZ388480 EINDL21TF
C 16	49.8	6.1	429	10	CG803197	CG803197 1118039H0
C 17	49.6	6.1	753	5	BW073695	BW073695 BW073695
C 18	49.6	6.1	791	5	BW143210	BW143210 BW143210
C 19	49.6	6.1	930	9	AZ549806	AZ549806 ENTFL24TR
C 20	49.4	6.1	1068	10	CNS00J4X	AL075933 Drosophil
C 21	49.2	6.1	641	9	AQ946120	AQ946120 Sheared D
C 22	49.2	6.1	1101	10	CNS00FMC	AL070972 Drosophil

23	49.2	6.1	1131	11	CNS034FO	AL227373 Tetraodon
C 24	49	6.0	878	10	CNS0187R	AL108993 Drosophil
C 25	49	6.0	1101	10	CNS0176D	AL107647 Drosophil
C 26	48.6	6.0	891	10	CW952279	CW952279 TcB38.1.B
C 27	48.4	6.0	1101	10	CNS00807	AL069440 Drosophil
C 28	48.2	5.9	1183	1	AJ926017	AJ926017 Drosophil
C 29	48	5.9	811	8	CK092809	CK092809 EHAF16TR
C 30	47.8	5.9	735	10	CW960936	CW960936 AIAA-aacl
C 31	47.8	5.9	767	10	EX193880	EX193880 Danio rer
C 32	47.6	5.9	332	5	BW139522	BW139522 BW139522
C 33	47.6	5.9	739	5	BW257592	BW257592 BW257592
C 34	47.6	5.9	783	5	BW131201	BW131201 BW131201
C 35	47.6	5.9	785	5	BW127705	BW127705 BW127705
C 36	47.4	5.8	320	1	AV984612	AV984612 AV984612
C 37	47.4	5.8	352	5	BW269652	BW269652 BW269652
C 38	47.4	5.8	354	5	BW269352	BW269352 BW269352
C 39	47.4	5.8	478	5	BW262060	BW262060 BW262060
C 40	47.4	5.8	496	3	BP018342	BP018342 BP018342
C 41	47.4	5.8	575	5	BW430026	BW430026 BW430026
C 42	47.4	5.8	713	5	BW480796	BW480796 BW480796
C 43	47.4	5.8	987	10	CNS014PQ	AL104456 Drosophil
C 44	47.2	5.8	580	11	CR479976	CR479976 mch2-190L
C 45	47.2	5.8	685	10	CG957658	CG957658 MBEE662TR

ALIGNMENTS

RESULT 1  
CNS0039G/c

LOCUS  
DEFINITION  
Drosophila melanogaster genome survey sequence TET3 end of BAC #  
BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
ACCESSION  
AL063921  
VERSION  
AL063921.1 GI:4941778  
KEYWORDS  
GSS.  
SOURCE  
Drosophila melanogaster (fruit fly)  
ORGANISM  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE  
1 (bases 1 to 1101)  
Genoscope.  
AUTHORS  
Direct Submission  
TITLE  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
JOURNAL  
- Web : www.genoscope.cns.fr)  
COMMENT  
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Oseegawa and Aaron Mamoer in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's Pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

FEATURES  
source

1. .1101  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACR08K10"  
/clone\_lib="RPCI-98"  
/note="end : TET3"

ORIGIN

```
Query Match      7.7%; Score 62.8; DB 10; Length 1101;
Best Local Similarity 15.8%; Pred. No. 0.00051;
Matches 102; Conservative 272; Mismatches 270; Indels 0; Gaps 0;

QY 11 AAAGATTGTTAGAGTTCACCTGAGTACATTTCTGTTTCTTCTATCTGATGGTCT 70
DB 1061 DADRWAADDGAGTWTATWWWWWWATWTDWWDKWWWWATAAATCTATWRTWARWD 1002

QY 71 TTTCTGATCAAAATTTCTGAAGGAGGAGGAGCTTTATATAGTAGTCAATAAAG 130
DB 1001 AGRDRGAGRRDRDAATDADGAGRRDGRKOKKDRKDDGDDKGGKKKKAATAKAT 942

QY 131 TTGGTATTCCTCAATTTTAGTAATTTTTCAGTCAAGAAACAATTCCTGGTATTACAAA 190
DB 941 WDDWDWDKWDGWDGAKDRKADDDGAGDKDDGKDDADDGTDGTDKDDDDKDDKDDW 882

QY 191 AGATTTTGGTGGTTCGTAAGTCTGAGATAAATACTCACAGCAATTTTACACGAT 250
DB 881 AKGTGWDATWAAATDWWWWGADADWTTWDAADDDWADDDWDAWKKWDDAWAWGART 822

QY 251 CATATGACCTACTTATGACAGCAGTTTTCGAGGTTTTCAGGTTTATGATGATATG 310
DB 821 ADRRDWDGDRAGRGSGARKRRDRKRDADDDDDAATTTTTRDTRDTRDTRDTRD 762

QY 311 TTAATGACTTTAGGGTAGAATTTGAAGGTTCTTATGAGAAATTTTGAACCTGAAAGCA 370
DB 761 TWTWRAADRTWDRDDDDDRDRAGTAGRKWRRTWKRWRKRDRTWDDADDDTARDRR 702

QY 371 GGTACCTGAGATAGCCAAAGCTACAAATTTTTCCTGTTGCTCGAAATGCTACAAAT 430
DB 701 GDDGADAGKGTGRRRRDRATWDRDADWADAATTTTDTDDWDRRRKRGARR 642

QY 431 GTGATAAAGTTTATAGTACTAGAGAAATAACGGCGTTGTTGACAGTCTCTTAATGTA 490
DB 641 RRTTARAADWMTWKAWDQKWDKTRADRWDRWAADTTWTDARKADRWDAKARWRARR 582

QY 491 ATGTTGTTATGATATGCTAGTGGTAGTATTCCTTTAGCACCTTATATGCTGGTGG 550
DB 581 DRARAARDRWTTGKTTTATTTWAARAAMWAWWATTTATTTTWTWTWTWTWT 522

QY 551 TTGGTGCAGATATATAAAGTTTTTATAGGTATATCATGCTGCTTCTTATCAAGTTA 610
DB 521 TTTTWTWTTAAAWAAWWTATWAAATATAAATAAATAAATAAATTTTTTTTTTAAW 462

QY 611 AGTTGGTGCACACTACCTCAATGTTAATGTTAATGTTGTTT 654
DB 461 TAWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 418
```

```
RESULT 2
CNS0100X      1101 bp      DNA      linear      GSS 26-JUL-1999
LOCUS      Drosophila melanogaster genome survey sequence sp6 end of BAC
DEFINITION      BACN03G04 of DrosBAC library from Drosophila melanogaster (fruit
                  fly), genomic survey sequence.
ACCESSION      AL098379
VERSION      AL098379.1 GI:5609990
KEYWORDS      Drosophila melanogaster (fruit fly)
SOURCE      Drosophila melanogaster
ORGANISM      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
              Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
              Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 1101)
AUTHORS      Genoscope.
TITLE      Direct Submission
JOURNAL      Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
              BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
              - Web : www.genoscope.cns.fr)
COMMENT      Determination of this BAC-end sequence was carried out as part of a
              collaboration with the European Drosophila Genome Project (EDGP) -
              http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC
              library (Dros BAC) was made by Alain Billaud at CEPH (Centre
```

```
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBel0BAC11.

FEATURES             Location/Qualifiers
     source            1..1101
     /organism="Drosophila melanogaster"
     /mol_type="genomic DNA"
     /db_xref="taxon:7227"
     /clone="BACN03G04"
     /clone_lib="DrosBAC"
     /plasmid="pBel0BAC11"
     /note="end : sp6"
```

## ORIGIN

```
Query Match      7.5%; Score 60.8; DB 10; Length 1101;
Best Local Similarity 17.6%; Pred. No. 0.0015;
Matches 124; Conservative 275; Mismatches 306; Indels 1; Gaps 1;

QY 43 TTTGTTTTTCTTATCTGATGGTCTTTTCTGATGCAAAATTTTCTCAAGGGAGGGA 102
DB 389 KTTTTRTTKGGGTGKRTGNGTGTGAGKDGKDGKGAATTTTDTTDTTDTTDTW 448

QY 103 GGACTTTATATAGGTAGTCAATATAAGTTGATATCCCAATTTTAGTAATTTTTCAG 162
DB 449 DTDWTDWTDWTTTWTWTTTGTGKKWDTKTKWKKDDGTTWTDWTDWTDWTDWTD 508

QY 163 GAAGAAACAATTCCTGGTATTACAAAAGATTTTTCGTTAGTCTTGAAGTCTGAG 222
DB 509 DVWADNAGWTTDKDKTKAKKRTTDTTKTKTGTGKTKDRGMRKTDTWGWGDR 568

QY 223 ATAAATACTCACAGCAATTTTACAGCATATAGCCCTACTTATGCAAGCAATTTTGA 282
DB 569 DGARADRAANAKDDKADGADTADTKTDGWTGTTTADTDDTDADWWDKWDKAD 628

QY 283 GGGTTTAGTGGTATCATTTGGATATTTATGTAATGACTTTTAGGGTGAATTTTCA 342
DB 629 AGRKGGWGWGKGGKWTTKTKKKTDTKTCTDTKWKWKDKWTDKDKRGGDGRKDK 688

QY 343 TATGAGAATTTTGAACCTGAAAGCAATGGTACCCCTGAGAAATAGCCAAAGCTTACA 402
DB 689 GTGGDGSKAKGDRAWAWRAKATRAAAADATAAATAAKATAWATKTTTCTTTT 748

QY 403 TTTGCTTGTCTCGAAATGCTACAATAGTATGATATAGTTTATAGTACTAGAGATA 462
DB 749 TTTTGTGTTTATKADRDDDDKAKRDKDKDGWDGDDKDWKAG-RDRDDDTTRK 807

QY 463 GCGGTTGTGACAGCTCTTAATGTAATGTTTCTTATGATATTTGCTAGTGTAGTAT 522
DB 808 DRAGKKKAKKKHKTAMWADDTGTANATWAKTATATKAKATKDKBRGGRGGGRGD 867

QY 523 CTTTATGACCTTATATGCTGCTGGTGTGTCAGATTTATATAAGTTTATAGGTATA 582
DB 868 RGAGRRKARGRKRRRRKDKRKRKRTKAGDGTGAGTDKAGDKKADGSGWGWKRW 927

QY 583 TCATTGCTTAAGTTTCTTATCAAGTAAAGTTTGTGTCAACTCCCTCTTAATGTTA 642
DB 928 DSKTWKTTWTGTTTDTTWTGRRNRGRGAGRRRGRKAGKAGGAGGKAWWADABWT 987

QY 643 ACTATGTTGTTGGTGGGGTTTATTACCATAGGTTTGTAGGTGATAGGCATGAGAGTA 702
DB 988 WTDWWDKWKWTTDWGKGGWGDGKGGKADGAKRNRKNRDGMRYGNRAWKWMDAA 1047

QY 703 GAATAGCTTACCTACTCTATGATATTTATGAGTCTCTCTAGAACTA 748
DB 1048 KARWADDWDADYDDDDADADYWKWKAATWMBWCDCTAKWTANWKNOW 1093
```

## RESULT 3

```
EX173672/c      829 bp      DNA      linear      GSS 13-MAR-2003
LOCUS      BX173672
DEFINITION      Danio rerio genomic clone DKEY-150M6, genomic survey sequence.
ACCESSION      BX173672
```

## RESULT 4







collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammeter in Piter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

```
FEATURES
    source          Location/Qualifiers
    1..1101
        /organism="Drosophila melanogaster"
        /mol_type="genomic DNA"
        /db_xref="taxon:7227"
        /clone="BACR48P19"
        /clone_lib="RPCI-98"
        /note="end : TET3"

ORIGIN
    Query Match      6.5%; Score 53; DB 10; Length 1101;
    Best Local Similarity 21.7%; Pred. No. 0.1;
    Matches 123; Conservative 182; Mismatches 259; Indels 2; Gaps 1;

QY 125 ATAAAGTGGTATCCCAATTTTGTAGTAATTTTTCAGCTGAAGAAACAATTCCTCGGTATTA 184
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 476 AAAAAAATAAAATTTTWTWKATTTTWTDAADDKAAAWDTTWTATTTTATTTTATTTT 535
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 185 CAAAAAAGATTTTGGTGTAGGCTTGTAGTAAGTCTCGAGATAAATATCTACAGCAATTTTA 244
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 536 ATTAATATTTTWTWKATKTDTWKAKTAATAAATAAATTTTATTTTATTTTATTTT 595
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 245 CAGGATCATATACCCCTACTTATGCAAGCAGTGTTCGAGGTTTGTAGTGTATCATTTGAT 304
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 596 WTTAKTWAQKWADATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 655
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 305 ATATGTTAATGACTTTAGGTTAGAAATTTGAAGGTTCTTATGAGAAATTTGAACCTGAAA 364
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 656 TTTTAKTNTTTTATTAATAAATAAATTTTDTWAAAWTWTTKKKKKKAAADKKWD 715
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 365 GACAATGGTACCTGAGATAGCAAGCTACAAATTTTGTCTTGTCTCGAAATGCTA 424
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 716 AKWDGAKKATKTKKDKKAA--WAAADKKKRGKGGKGGKGGKGGKGGKGGKGGKGG 773
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 425 CAAATAGTGATAATAAGTTTATAGTACTAGAGAAATAACGCGGTTGTTGCAAGTCTCTTA 484
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 774 GWGKGGKAGDDADKDKTKKKKKAATTTTKKKKGGKGGKGGKGGKGGKGGKGGKGGKGG 833
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 485 ATGTAATGTTGTTATGATATTCGTAGTGGTAGTATTCCTTTAGCACCTTATATATGTGTG 544
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 834 AAAAAAAKTKDKGKKKKTKKTKTKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 893
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 545 CTGGTGTGTCAGATATATAAGTTTATAGGTTATATCATATCATGTCCTTAATTTCTTATC 604
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 894 GKGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGG 953
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 605 AAGTTAAGTTTGGTGTCACTACCTCTCAATGTTAATGTTAATGTTTGTGGGGGTT 664
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 954 KKKKKKADAAGKTKKKAADAADAADWDATKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 1013
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 665 ATTACCATAGGTTGTAGGTAGG 690
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1014 KKKKKKTKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 1039
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 9
CNS010DV/c      976 bp   DNA   linear   GSS 26-JUL-1999
LOCUS           Drosophila melanogaster genome survey sequence SP6 end of BAC
DEFINITION
```

```
ACCESSION      BACN03N21 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
VERSION        AL098845
KEYWORDS       AL098845.1 GI:5610456
SOURCE         Drosophila melanogaster (fruit fly)
ORGANISM       Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 976)
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CRPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.

FEATURES
    source          Location/Qualifiers
    1..976
        /organism="Drosophila melanogaster"
        /mol_type="genomic DNA"
        /db_xref="taxon:7227"
        /clone="BACN03N21"
        /clone_lib="DrosBAC"
        /plasmid="pBelobAC11"
        /note="end : SP6"

ORIGIN
    Query Match      6.5%; Score 52.6; DB 10; Length 976;
    Best Local Similarity 29.5%; Pred. No. 0.13;
    Matches 94; Conservative 92; Mismatches 132; Indels 1; Gaps 1;

QY 16 TTTGTTGAGGTGTACGCTGAGTACATTTGTTTTTTCTTCTATCTGATGGCTTTTCT 75
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 953 DMKTKTKTKGTATWATRYDDRTAATAATGTGTTTWTWTATKRTGTGTTTAT 894
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 76 GATGCAAAATTTTCTGAAGGGAGGAGGACTTTATATAGTAGTACAGTATAAAGTTGCT 135
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 893 CKTGATGTTGDTGGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGG 834
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 136 ATTCCTCAATTTTGTAGTA-ATTTTTCAGCTGAAGAAACAAATTCCTGGTATTACAAAAAGAT 194
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 833 WGDTRRRRHAADKGAWAWAWWWTADAGDRWRTRAGNKTGTTKTRWGTTGAKRTDRGAK 774
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 195 TTTTCGGTTAGTCTTGTAGTCTGAGATAAATACTACAGCAATTTTACAGCATATA 254
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 773 AKRDTGGKRRGTATRWATRTATATATATTTATTTTAAHWWTATTTTATTTATTT 714
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 255 TGACCTCTACTATGCAAGCAGTTTTCGAGGGTTTGTAGTGTATCATTTGGATATATGTTAA 314
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 713 WGDWTWTATATKTTTGTAKATGAAAWTATTTATTTATTTATTTTAKATWTWT 654
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 315 TGACTTTAGGGTAGAATTT 333
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 653 ATWTATATWGTNWAATAT 635
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
BX152436
LOCUS           Danio rerio genomic clone DKEY-116F16, genomic survey sequence.
DEFINITION      Danio rerio
ACCESSION      BX152436
VERSION        BX152436.1 GI:27983930
KEYWORDS       GSS.
SOURCE         Danio rerio (zebrafish)
ORGANISM       Danio rerio
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Humphray, S.J., Huckle, E. and Durham, J.L.  
Direct Submission  
Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humphray@sanger.ac.uk Unpublished  
This sequence was generated from the SP6 end of BAC 116F16. 116F16 is part of the Daniokey BAC Library created by R. Plasterk and N.V. Keygene. Further details:  
http://www.sanger.ac.uk/Projects/D\_rerio/  
Location/Qualifiers

FEATURES  
source

1..630  
/organism="Danio rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/clone="DKEY-116F16"  
/tissue\_type="Testis"  
/note="vector pIndigoBAC-536"

ORIGIN

Query Match 6.3%; Score 51.2; DB 10; Length 630;  
Best Local Similarity 48.9%; Pred. No. 0.26; Indels 0; Gaps 0;  
Matches 137; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

Qy 418 AATGCTCAAAATAGTGAATAAAGTTATATAGTACTAGAGATAACGGCGTTGTGACAAG 477  
Db 134 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 193  
Qy 478 TCTCTTAATGAATGTTGTTGATATATCTAGTGTAGTATCTCTTAGACACCTTAT 537  
Db 194 ATTATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 253  
Qy 538 ATGTGTGCTGCTGTTGGTCAGATTATATAAGTTTATAGTATATCATTCGCCAAGTTT 597  
Db 254 ATTGTTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 313  
Qy 598 TCTTATCAAGTTAAGTTTGGTGTCAACTACCTCTAAATGTTAATACTATCTGTTGGT 657  
Db 314 TATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 373  
Qy 658 GGGGGTTATTACCAATAAGTTGTAGGTGATGGCATGAGA 697  
Db 374 ATTACTATTATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 413

RESULT 11  
BX139987  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Danio rerio (zebrafish)  
Danio rerio  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Humphray, S.J., Huckle, E. and Durham, J.L.  
Direct Submission  
Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humphray@sanger.ac.uk Unpublished  
This sequence was generated from the T7 end of BAC 99E7. 99E7 is part of the Daniokey BAC Library created by R. Plasterk and N.V. Keygene. Further details:  
http://www.sanger.ac.uk/Projects/D\_rerio/  
Location/Qualifiers

FEATURES  
source

1..844  
/organism="Danio rerio"

/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/clone="DKEY-99E7"  
/tissue\_type="Testis"  
/note="vector pIndigoBAC-536"

ORIGIN

Query Match 6.3%; Score 51; DB 10; Length 844;  
Best Local Similarity 45.3%; Pred. No. 0.29; Indels 1;  
Matches 229; Conservative 0; Mismatches 270; Indels 6; Gaps 1;

Qy 15 ATTGTGTTAGGTTTACGCTGAGTACATTTGTTTTTTTCTTATCTGATGGTCTTTTC 74  
Db 304 AATTATATTGTCAGTCATTTATTCCTTTTATTGATCATTAATAATAATAATAATGA 363  
Qy 75 TGATGCAAAATTTTCTGAAGGGAGAGAGACTTTATATAGGTAGTCAGTATAAAGTTGG 134  
Db 364 TAATAATAATGATAATAATAATAATGATGATAATAATAATAATAATAATAATAATAA 423  
Qy 135 TATTCCCAATTTTAGTAATTTTTCAGCTGAAGAACAAATTCCTGGTATTACAAAAAGAT 194  
Db 424 TAATAATAATAATGATAATAATAATAATAATATTGATAATAATAATAATAATAATGA 483  
Qy 195 TTTTGGCTTAGGTTCTTGATAAAGTCTGAGATAAAATCTCACAGCAATTTTACCGATCAT 254  
Db 484 TAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 543  
Qy 255 TGACCTTACTTATGCAAGCAGTTTTCGAGGGTTTAGTGGTATCATTTGGATATTATGTTAA 314  
Db 544 TGATAATAATAATGATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 603  
Qy 315 TGACTTTAGGTTAGAAATTTGAAGTTCTTATGAGAAATTTTGAACCTGAAAGACATGTA 374  
Db 604 TGATGATA-----ATAATGATAATAATAACAATAATAATAATAATAATAATAATAA 657  
Qy 375 CCCTGAGATAGCAAAAGCTACAAATTTTTCCTGCTCGAAATCTCTCAAAATAGTGA 434  
Db 658 TAATGATGATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATGA 717  
Qy 435 TAATAAGTTTATAGTACTAGAGATAACGGCGTTGTTGACAAGTCTCTTAAATGTAATGT 494  
Db 718 CAATAATAATAATGACAATAATAATAATAATAATAATAATAATAATAATAATAATAATGA 777  
Qy 495 TTGTTATGATATTGCTAGTGGTAGT 519  
Db 778 TACTAATAATAATGATAATAATAATAAT 802

RESULT 12

CW973509  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Ancylostoma caninum (dog hookworm)  
Ancylostoma caninum  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida; Ancylostomatoidea; Ancylostomatidae; Ancylostomatinae; Ancylostoma.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Mitrevva, M., McCarter, J.P., Pape, D., Ritter, E., Taagareishvili, R., Ronko, I., Martin, J., Wyllie, T., Dantes, M., Meyer, R., Messina, D., Waterston, R.H., Clifton, S.W. and Wilson, R.  
Genome Survey sequences from the parasitic nematode Ancylostoma caninum  
Unpublished (2004)  
Contact: Mitreva M  
Washington University in St. Louis  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: [nematode@watson.wustl.edu](mailto:nematode@watson.wustl.edu)  
Genomic DNA provided by John Hawdon ([mtcmjhb@uwmc.edu](mailto:mtcmjhb@uwmc.edu)) DNA  
sequenced by Washington University Genome Sequencing Center  
Class: shotgun.

[illegible]

RESULT 13	CNS012FZ	1094 bp	DNA	linear	GSS 26-JUL-1995
CNS012FZ/c					
LOCUS					
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC BACN07B02 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL101513				
VERSION	AL101513.1				
KEYWORDS	GSS.				

SOURCE	Drosophila melanogaster (fruit fly)			
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
REFERENCE	1 (bases 1 to 1094)			
AUTHORS	Genoscope.			
TITLE	Direct Submission			
JOURNAL	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage ; BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)			
COMMENT	- web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBAC11.			
FEATURES	Location/Qualifiers			
source	1..1094			
	/organism="Drosophila melanogaster"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:7227"			
	/clone="BACN07B02"			
	/clone_lib="DrosBAC"			
	/plasmid="pBelOBAC11"			
	/note="end : T7"			
ORIGIN				
Query Match	6.2%	Score 50.4;	DB 10;	Length 1094;
Best Local Similarity	33.1%	Pred. No. 0.42;		
Matches	224;	Conservative	93;	Mismatches 355; Indels 4; Gaps 1;
Qy	20	TTGTAGGTGTACGCAGTACATCTGTTTTTTCTTATCTGATGGTCTTTTCTGATG	79	
Db	1040	TTTTTTTGTGTTAKRWTTTAKWTTTTTTTADWRAAATTTTWTDRWTAATA	981	

## ORIGIN

Query Match	6.28;	Score 50.4;	DB 10;	Length 1094;
Best Local Similarity	33.1%;	Pred. No. 0.42;		
Matches	224;	Conservative 93;	Mismatches 355;	Indels 4; Gaps 1;
Qy	20	TTGTAGGTGTTACGCTGAGTACATTTGTTTTTTCTTATCTGATGGTCTTTTCTTGATG 79		
Db	1040	TTTTTTTGGWTKTAKRWTTTTTKAWTTTTTTTTTADWRRAAATTTTWDRWTAATA 981		
Qy	80	CAAAATTTTCTGAAGGAGGAGGAGCTTTATATAGTGTACGATATAAAAGTTGGTATTC 139		
Db	980	WATTAWTTTTTTWTGTGAAGRRGARAATTTAAATTTWAAAAAAMWWATWAAATWAAADAGR 921		
Qy	140	CCAATTTTAGTAATTTTTAGCTGGAAGAAACAATCTCGTGTATACAAAAGAAATTTTTCG 199		
Db	920	TKWATTTTAGGRTGRTGRTTTTTTAARGGRTWW---AAAWTTTTTWGATTTTTTWTATW 865		
Qy	200	CGTTAGTCTTGATAAGTCTGAGATAAACTCACAGCAATTTTACACGATCATATGACC 259		
Db	864	TGAAGAWTTTTAWTTTTTTTTTTTTTAAATTTWAATTTAAATTTATTTAAARWAWW 805		
Qy	260	CTACTTATGCAAGCAGTTTTTGCAGGTTTTAGTGGTATCATTTGGATATATATGTTAATGACT 319		
Db	804	TTTTTAWADTTTTTTTTTTRGRRRRTTGRWTTAAWTTTAAAAAGGTTTTTWTWTTTTTT 745		
Qy	320	TTAGGTGATTAATTTGAAGGTTCTTATGAGAAATTTTGAACCTGGAAGACAATGGTACCCCTG 379		
Db	744	WAAAAARAATTTTTTTTTTTTWWTTTTTTTTTTTAAATAATAAARTTTTTDGGGTG 685		
Qy	380	AGAATGCCAAAGCTACAAAATTTTTTGCTTTGTCTCGAAATGCTACAAATAGTGATAATA 439		
Db	684	RTTWRATATAAATTWAAKTTTTTTTGGRWARTTTTTTTRAAAAWMTGAWATGGRAGVTTTG 625		
Qy	440	AGTTTATAGTACTAGAGATAAAGCGGGCTGTTGACAAGTCTCTTAATGTAAATGTTTGGTT 499		
Db	624	RTTTGGRKAAATKTGTWAAATTTTTTGGTTTGGAAATTTAAGMRRTTTTTTTTTAAATTTTTTT 565		
Qy	500	ATGATATTGCTAGTGGTAGTATTCCTTTTACACCTTATATGTGTGCTGGTGGTGCAG 559		
Db	564	TGRAAGTTTTKGGGTGWTGKTBSGTTTTTAKWGAATTTTTTATKAAAAMWYKGGGAAAKK 505		
Qy	560	ATTATATAAAGTTTTTTAGGTATATCATTTGCTCAAGTTTTCTTATCAAGTTTAAAGTTGGTG 619		
Db	504	TTAATTAWTTTTTTTTTAAATWATKKTAGGAKGGTTTTTTTTTTTAKTKAKCGCTGAAGGKT 445		

Qy 620 TCAACTACCTTAAGTAACTACTATGTTGTTGGTGGGGTTATTACCATAGGTTG 679  
 Db 444 AATTCTWTTTSSSTSSGTTAAATAATGTTGTTTGGTTTGGGGGAAATGTGGGGSTT 385  
 Qy 680 TAGGTGATAGCATGA 695  
 Db 384 TTGGTTGGGGGATAA 369

RESULT 14  
 CG807194 471 bp DNA linear GSS 10-NOV-2003  
 LOCUS 1118078H01.y1 1118 - RescueMu Grid S Zea mays genomic, genomic survey sequence.

ACCESSION CG807194  
 VERSION CG807194.1 GI:38244754  
 KEYWORDS GSS.  
 SOURCE Zea mays

## ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 471)

Walbot,V.

Maize genomic sequences found using engineered RescueMu transposon

Unpublished (2001)

Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Plate: 1118078 column: 2

Class: transposon-tagged.

Location/Qualifiers

1..471

/organism="Zea mays"

/mol\_type="genomic DNA"

/culivar="mixed background W23/A188/B73"

/db\_xref="taxon:4577"

/tissue\_type="leaf"

/lab\_host="DH10B"

/dev\_stage="adult"

/clone\_lib="1118 - RescueMu Grid S"

/note="Organ: leaf; Vector: RescueMu (engineered from pBlueScript backbone); Site 1: BamHI; Site 2: BglII;

RescueMu is a 4.9 kb, modified maize Mu transposon

designed to allow plasmid rescue from total genomic DNA.

Mu elements insert preferentially into transcription

units. For more information on RescueMu, go to the web

site 'www.zmldb.iastate.edu' and follow the links for

'RescueMu.' Grid S was grown at San Diego in 2002. DNA was

extracted from leaf strips, double digested using BamHI

and BglII, and ligated to form circular plasmids. DH10B

cells were transformed and then screened on LB plates with

ampicillin."

## ORIGIN

Query Match 6.2%; Score 50.2; DB 10; Length 471;  
 Best Local Similarity 45.1%; Pred. No. 0.43;  
 Matches 187; Conservative 0; Mismatches 228; Indels 0; Gaps 0;

Qy 108 TTATATAGTGTAGTATAGTAAAGTTGGTATCCCAATTTTGTAGTATTTTTCAGCTGAAGA 167

Db 1 TTATATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 60

Qy 168 AACAAATTCCTGGTATTACAAAAGAGATTTTTCGGTTAGGCTCTTGATAGTCTGAGATAAA 227

Db 61 TAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 120

Qy 228 TACTCACAGCAATTTTACACGATCATATGACCCCTACTTATGCAAGCAGCTTTTGCAGGTT 287

Db 121 TAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 180  
 Qy 288 TAGTGGTATCATTTGGATATATTATGTTAATGACTTTTAGGGTAGAATTTTGAAGGTTCTTATGA 347  
 Db 181 TAATTATAATTATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 240  
 Qy 348 GAATTTTGAACCTGAAGACAAATGGTACCTGAGATAGCCCAAGCTAGCAAAATTTTTCG 407  
 Db 241 TAATTATAATTATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 300  
 Qy 408 TTGTCTCGAAATGCTACAAATAGTATAGTATAGTTTATAGTACTAGAGAATAACGGCGT 467  
 Db 301 TATTATAATTATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 360  
 Qy 468 TGTTGACAAAGTCTCTTAAATGTTAAATGTTTGTATGATATTTGCTAGTGGTAGTATT 522  
 Db 361 TAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 415

## RESULT 15

BZ388480/c

LOCUS

DEFINITION

BZ388480 670 bp DNA linear GSS 30-APR-2003

genomic survey sequence.

ACCESSION BZ388480

VERSION BZ388480.1 GI:302335017

KEYWORDS GSS.

SOURCE

Entamoeba invadens

ORGANISM

Entamoeba invadens

Eukaryota; Entamoebidae; Entamoeba.

REFERENCE

1 (bases 1 to 670)

Wang,Z., Samuelson,J., Clark,C.G., Eichinger,D., Paul,J., van

Dellen,K., Hall,N., Anderson,I. and Loftus,B.

Gene discovery in the Entamoeba invadens genome

Mol. Biochem. Parasitol. 129 (1), 23-31 (2003)

JOURNAL

PUBMED

12798503

COMMENT

Other\_GSSs: EINDL21TR

Contact: Brendan Loftus

Department of Eukaryotic Genomics

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-3543

Fax: 301-838-0208

Email: enta@tigr.org

DNA was provided by Daniel Eichinger

Seq primer: TP

Class: sheared ends.

Location/Qualifiers

1..670

/organism="Entamoeba invadens"

/mol\_type="genomic DNA"

/strain="IP-1"

/db\_xref="taxon:33085"

/clone="EINDL21"

/clone\_lib="E1\_10\_12 kb"

/notes="Vector: pHOS2; Site 1: BstXI; Total genomic DNA was

isolated from early log phase trophozoites of E. invadens

IP-1 using a Qiagen plant DNA extraction kit. A shotgun

medium-size plasmid library (average insert size of 10 -

12 kb) was generated by random mechanical shearing of E.

invadens genomic DNA, repairing the ends of DNA fragments

with T4 Polymerase, adding BstXI adaptors and ligating

into the BstXI site of a pUC-derived vector pHOS2."

ORIGIN

Query Match 6.2%; Score 50; DB 9; Length 670;

Best Local Similarity 43.8%; Pred. No. 0.49;

Matches 218; Conservative 0; Mismatches 280; Indels 0; Gaps 0;

Qy 117 TAGTCAGTAAAGTTGGTATTCCTCAATTTTGTAGTATTTTTCAGCTCAAGAAACAATCC 176

Db 665 TAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 606



Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	48.2	5.9	6113	6	US-10-240-708-14	Sequence 14, Appl
2	45.6	5.6	20500	6	US-10-829-826B-16	Sequence 16, Appl
3	45.6	5.6	27608	6	US-10-829-826B-14	Sequence 14, Appl
4	45.6	5.6	27608	6	US-10-829-826B-15	Sequence 15, Appl
5	45.6	5.6	27635	6	US-10-829-826B-17	Sequence 17, Appl
6	45.2	5.6	1044	7	US-11-066-648A-1	Sequence 1, Appl
C 7	41.4	5.1	1773	6	US-10-750-185-58930	Sequence 58930, A
C 8	41.4	5.1	139054	7	US-11-121-086-96	Sequence 96, Appl
9	41.2	5.1	31028	6	US-10-829-826B-21	Sequence 21, Appl
10	41.2	5.1	31028	6	US-10-829-826B-22	Sequence 22, Appl
11	41.2	5.1	31028	6	US-10-829-826B-26	Sequence 26, Appl
12	41.2	5.1	31032	6	US-10-829-826B-23	Sequence 23, Appl
13	41.2	5.1	31100	6	US-10-829-826B-24	Sequence 24, Appl
14	40	4.9	6317	6	US-10-240-708-11	Sequence 11, Appl
15	39.2	4.8	1289	6	US-10-750-185-39647	Sequence 39647, A
16	39.2	4.8	6556	6	US-10-240-708-76	Sequence 76, Appl
17	39.2	4.8	27733	6	US-10-829-826B-18	Sequence 18, Appl
18	38.4	4.7	96388	7	US-11-117-187-196	Sequence 196, Appl
19	38	4.7	645179	6	US-10-595-561-13293	Sequence 13293, A
20	37.8	4.6	7664	6	US-10-240-708-83	Sequence 83, Appl
C 21	37.6	4.6	159781	7	US-11-121-086-92	Sequence 92, Appl
C 22	37.6	4.6	200628	7	US-11-121-086-62	Sequence 62, Appl
C 23	37.6	4.6	207908	7	US-11-112-908-21	Sequence 21, Appl

Db 2806 TTATTGGAATTAAAGTAATAGATAAATTTTATTATTATTAGAAATTTTTTAAATTTTTTAG 2865

Qy 302 GATATTATGTTAATGACTTTTAGGGTAGAATTTGAAAGGTTCTTATG-----AGAAATTTTGAA 357

Db 2866 TAGTTATTATTTTATATTTTGTGTAATTAATGATGTTTATGCGGTTTATAGTTTGT 2925

Qy 358 CCTGAAAGCAATGGGTACCCGTAGAGATAGCCAAAGCTACAAATTTTTTGTCTCGA 417

Db 2926 TTTTATAGATATATATAAATGGAATATATAGGAATTTAGATTTTGTAGTTTGTCT 2985

Qy 418 AATGCTACAAATAGTGATATAAGTTTATAGTACTAGAGAATAACCGCGTTGTCACAAG 477

Db 2986 TTTATATAGTATAATGTTAGTATGAGATTTATTTATGTTAGTAGTTGTTTTTTTTTATG 3045

Qy 478 TCTCTTAATGTAATGTTGTTATGATATTCGTAGTGGTAGTATTCCTTTAGCACCTTAT 537

Db 3046 TTTAGTAGTATTTTATTTGATGGGTTGTTAGTTTATTTGTTTATAGTTGAAGGATAT 3105

Qy 538 ATGTGCTGCTGTTGTTGTCAGATTTATATAAGTTTTTTAGGTATATCATTTGCCCTAAGTTT 597

Db 3106 AGGATGTTTTTATGTTTGTGATGTCATPAAAGTTTTTATAATATTTTATATAGGAT 3165

Qy 598 TCTTATCAAGTTAAGTTTGGTGT 620

Db 3166 TGTGTGAACGTAAGTTTATAT 3188

RESULT 2

US-10-829-826B-16

; Sequence 16, Application US/10829826B

; Publication No. US20050266397A1

; GENERAL INFORMATION:

; APPLICANT: Ecker, David J.

; APPLICANT: Hofstadler, Steven A.

; APPLICANT: Sampath, Rangarajan

; APPLICANT: Blyn, Lawrence B.

; APPLICANT: Hall, Thomas A.

; APPLICANT: Massire, Christian

; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES

; FILE REFERENCE: IBIS0075-100 (DIBIS-0058US)

; CURRENT APPLICATION NUMBER: US/10/829, 826B

; CURRENT FILING DATE: 2004-04-22

; NUMBER OF SEQ ID NOS: 105

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 16

; LENGTH: 20500

; TYPE: DNA

; ORGANISM: Coronavirus

US-10-829-826B-16

Query Match 5.6%; Score 45.6; DB 6; Length 20500;

Best Local Similarity 49.2%; Pred. No. 0.38;

Matches 120; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

Qy 309 TGTTAATGACTTTAGGGTAGAATTTGAAAGGTTCTTATGAGAATTTTGAACCTGAAAGACA 368

Db 9378 TGGAACTGACCTAATGGTGAAATCTATGTTGTTATGTTGATGAAGAGTTGCACAAAG 9437

Qy 369 ATGGTACCTGAGAATAGCCAAAGCTACAAATTTTTTGTCTCGAAATGCTACAAA 428

Db 9438 AGTGCCACCAGATAATTTAGTTACTAACAAATATTTAGCATGGCTCTATGCGGCAATTAT 9497

Qy 429 TAGTGATAATAAGTTTATAGTACTAGAGAATAACGGGTTGTTGACAAAGTCTCTTAATCT 488

Db 9498 TAGTTTAAAGGAGAGTAGTTTCTCGCTGCCTAAATGTTGGAGAGTACTACTGTAGTGT 9557

Qy 489 AAATGTTTGTATGATATTTGCTAGTGGTAGTATTCCTTTAGCACCTTATATGTGTGCTGG 548

Db 9558 TGATGATTATAAAGTGGCTGGTGACAATGTTTTTACACCAATTTTCTACTAGTACCGC 9617

Qy 549 TGTT 552

Db 9618 TATT 9621

RESULT 3

US-10-829-826B-14

; Sequence 14, Application US/10829826B

; Publication No. US20050266397A1

; GENERAL INFORMATION:

; APPLICANT: Ecker, David J.

; APPLICANT: Hofstadler, Steven A.

; APPLICANT: Sampath, Rangarajan

; APPLICANT: Blyn, Lawrence B.

; APPLICANT: Hall, Thomas A.

; APPLICANT: Massire, Christian

; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES

; FILE REFERENCE: IBIS0075-100 (DIBIS-0058US)

; CURRENT APPLICATION NUMBER: US/10/829, 826B

; CURRENT FILING DATE: 2004-04-22

; NUMBER OF SEQ ID NOS: 105

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 14

; LENGTH: 27608

; TYPE: DNA

; ORGANISM: Coronavirus

US-10-829-826B-14

Query Match 5.6%; Score 45.6; DB 6; Length 27608;

Best Local Similarity 49.2%; Pred. No. 0.42;

Matches 120; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

Qy 309 TGTTAATGACTTTAGGGTAGAATTTGAAAGGTTCTTATGAGAATTTTGAACCTGAAAGACA 368

Db 9378 TGGAACTGACCTAATGGTGAAATCTATGTTGTTATGTTGATGAAGAGTTGCACAAAG 9437

Qy 369 ATGGTACCTGAGAATAGCCAAAGCTACAAATTTTTTGTCTCGAAATGCTACAAA 428

Db 9438 AGTGCCACCAGATAATTTAGTTACTAACAAATATTTAGCATGGCTCTATGCGGCAATTAT 9497

Qy 429 TAGTGATAATAAGTTTATAGTACTAGAGAATAACGGGTTGTTGACAAAGTCTCTTAATCT 488

Db 9498 TAGTTTAAAGGAGAGTAGTTTCTCGCTGCCTAAATGTTGGAGAGTACTACTGTAGTGT 9557

Qy 489 AAATGTTTGTATGATATTTGCTAGTGGTAGTATTCCTTTAGCACCTTATATGTGTGCTGG 548

Db 9558 TGATGATTATAAAGTGGCTGGTGACAATGTTTTTACACCAATTTTCTACTAGTACCGC 9617

Qy 549 TGTT 552

Db 9618 TATT 9621

RESULT 4

US-10-829-826B-15

; Sequence 15, Application US/10829826B

; Publication No. US20050266397A1

; GENERAL INFORMATION:

; APPLICANT: Ecker, David J.

; APPLICANT: Hofstadler, Steven A.

; APPLICANT: Sampath, Rangarajan

; APPLICANT: Blyn, Lawrence B.

; APPLICANT: Hall, Thomas A.

; APPLICANT: Massire, Christian

; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES

; FILE REFERENCE: IBIS0075-100 (DIBIS-0058US)

; CURRENT APPLICATION NUMBER: US/10/829, 826B

; CURRENT FILING DATE: 2004-04-22

; NUMBER OF SEQ ID NOS: 105

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 15

; LENGTH: 27608

; TYPE: DNA

; ORGANISM: Coronavirus

US-10-829-826B-15



Query Match	5.6%	Score 45.6	DB 6	Length 27608
Best Local Similarity	49.2%	Pred. No. 0.42		
Matches 120	Conservative 0	Mismatches 124	Indels 0	Gaps 0
Qy	309	TGTTAATGACTTTAGGGTAGAATTTGAAAGGTTCTTATGAGAATTTTGAACCTGAAAGACA	368	
Db	9378	TGGAACCTGACCTAATGGGTGAAATTTCTATGGTGGTTATGTTGATGAAGAGGTTGCACAAG	9437	
Qy	369	ATGTCACCTCGAAGATAGCCAAAGTACAAATTTTTTGGCTTGTCTCGAAATGCTACAAA	428	
Db	9438	AGTCCACCAAGATAATTTAGTTACTTAACAATAATGTAGCAATGGCTCTATCGGCAATTAT	9497	
Qy	429	TAGTGATAATTAAGTTTATAGTACTACAGAGAATAACGGCGTTTGTTCACAAGTCTCTTAATGT	488	
Db	9498	TAGTGTTAAGAGAGTAGTTTCTCGTCGCCTAAATGGTTGGAGAGTACTACTGTTAGTGT	9557	
Qy	489	AAATGTTTGTATGATATTCGTAGTGGTAGTATTCCTTTTAGCACCTTTATATGTGTCGTG	548	
Db	9558	TGATGATTAAATAAGTGGGCTGGTGACAATGGTTTTACACCAATTTTCTACTAGTACCGC	9617	
Qy	549	TGTT 552		
Db	9618	TATT 9621		

```

RESULT 5
US-10-829-826B-17
; Sequence 17, Application US/10829826B
; Publication No. US20050266397A1
; GENERAL INFORMATION:
; APPLICANT: Ecker, David J.
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Blyn, Lawrence B.
; APPLICANT: Hall, Thomas A.
; APPLICANT: Messire, Christian
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES
; FILE REFERENCE: IBIS0075-100 (DIBIS-00580US)
; CURRENT APPLICATION NUMBER: US/10/829.826B
; CURRENT FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 27635
; TYPE: DNA
; ORGANISM: Coronavirus
US-10-829-826B-17

```

	Query Match	5.6%	Score 45.6;	DB 6;	Length 27635;																																																																																																
	Best Local Similarity	49.2%;	Pred. No. 0.42;																																																																																																		
	Matches 120;	Conservative 0;	Mismatches 124;	Indels 0;	Gaps 0;																																																																																																
Qy	309	TGTTAATGACHTTTAGGGTAGAATTTGAAGGTTCTTATGAGAAATTTTGAACCTGAAAGACA	368																																																																																																		
						Qy	369	ATGCTACCTCGAGATACCCAAAGCTACAAATTTTTTTCCTTGTCTCGAAATGCTACAAA	428									Db	9438	AGTGCCACCAAGATAATTTAGTATTCTAACAAATATTTGTAGCATGGGCTCTATGCCGCANTTAT	9497									Qy	429	TAGTGATAATAAGTTTATTAGTACTAGAGAATAACGGCGTTCTTGACAAAGTCTCTTAATGT	488									Db	9498	TAGTGTTTAAGAGAGTAGTATTTCTCGCTGCCTTAATGGTTTGAGAGTACTACTGTTAGTGT	9557									Qy	489	AAATGTTTTGTTTATGATATTCGTAGTGGTAGTATTCCTTTAGCACCTTATATGTGTGCTGG	548									Db	9558	TGATGATTTATAATAAGTGGGCTGGTGACAAATGGTTTTACACCAATTTCTCTAGTACCGC	9617									Qy	549	TGTTT	552									Db	9618	TATT	9621								
Qy	369	ATGCTACCTCGAGATACCCAAAGCTACAAATTTTTTTCCTTGTCTCGAAATGCTACAAA	428									Db	9438	AGTGCCACCAAGATAATTTAGTATTCTAACAAATATTTGTAGCATGGGCTCTATGCCGCANTTAT	9497									Qy	429	TAGTGATAATAAGTTTATTAGTACTAGAGAATAACGGCGTTCTTGACAAAGTCTCTTAATGT	488									Db	9498	TAGTGTTTAAGAGAGTAGTATTTCTCGCTGCCTTAATGGTTTGAGAGTACTACTGTTAGTGT	9557									Qy	489	AAATGTTTTGTTTATGATATTCGTAGTGGTAGTATTCCTTTAGCACCTTATATGTGTGCTGG	548									Db	9558	TGATGATTTATAATAAGTGGGCTGGTGACAAATGGTTTTACACCAATTTCTCTAGTACCGC	9617									Qy	549	TGTTT	552									Db	9618	TATT	9621														
						Db	9438	AGTGCCACCAAGATAATTTAGTATTCTAACAAATATTTGTAGCATGGGCTCTATGCCGCANTTAT	9497									Qy	429	TAGTGATAATAAGTTTATTAGTACTAGAGAATAACGGCGTTCTTGACAAAGTCTCTTAATGT	488									Db	9498	TAGTGTTTAAGAGAGTAGTATTTCTCGCTGCCTTAATGGTTTGAGAGTACTACTGTTAGTGT	9557									Qy	489	AAATGTTTTGTTTATGATATTCGTAGTGGTAGTATTCCTTTAGCACCTTATATGTGTGCTGG	548									Db	9558	TGATGATTTATAATAAGTGGGCTGGTGACAAATGGTTTTACACCAATTTCTCTAGTACCGC	9617									Qy	549	TGTTT	552									Db	9618	TATT	9621																				
						Qy	429	TAGTGATAATAAGTTTATTAGTACTAGAGAATAACGGCGTTCTTGACAAAGTCTCTTAATGT	488									Db	9498	TAGTGTTTAAGAGAGTAGTATTTCTCGCTGCCTTAATGGTTTGAGAGTACTACTGTTAGTGT	9557									Qy	489	AAATGTTTTGTTTATGATATTCGTAGTGGTAGTATTCCTTTAGCACCTTATATGTGTGCTGG	548									Db	9558	TGATGATTTATAATAAGTGGGCTGGTGACAAATGGTTTTACACCAATTTCTCTAGTACCGC	9617									Qy	549	TGTTT	552									Db	9618	TATT	9621																																
						Db	9498	TAGTGTTTAAGAGAGTAGTATTTCTCGCTGCCTTAATGGTTTGAGAGTACTACTGTTAGTGT	9557									Qy	489	AAATGTTTTGTTTATGATATTCGTAGTGGTAGTATTCCTTTAGCACCTTATATGTGTGCTGG	548									Db	9558	TGATGATTTATAATAAGTGGGCTGGTGACAAATGGTTTTACACCAATTTCTCTAGTACCGC	9617									Qy	549	TGTTT	552									Db	9618	TATT	9621																																												
						Qy	489	AAATGTTTTGTTTATGATATTCGTAGTGGTAGTATTCCTTTAGCACCTTATATGTGTGCTGG	548									Db	9558	TGATGATTTATAATAAGTGGGCTGGTGACAAATGGTTTTACACCAATTTCTCTAGTACCGC	9617									Qy	549	TGTTT	552									Db	9618	TATT	9621																																																								
						Db	9558	TGATGATTTATAATAAGTGGGCTGGTGACAAATGGTTTTACACCAATTTCTCTAGTACCGC	9617									Qy	549	TGTTT	552									Db	9618	TATT	9621																																																																				
						Qy	549	TGTTT	552									Db	9618	TATT	9621																																																																																
						Db	9618	TATT	9621																																																																																												

```

RESULT 6
US-11-066-648A-1
; Sequence 1, Application US/11066648A
; Publication No. US20050260621A1
; GENERAL INFORMATION:
; APPLICANT: MCBRIDE, JERE W.
; APPLICANT: WALKER, DAVID H.
; APPLICANT: DOYLE, CHRISTOPHER KUYLER
; TITLE OF INVENTION: AN IMMUNOREACTIVE 38-KDA FERRIC BINDING PROTEIN OF
; TITLE OF INVENTION: EHRLICHIA CANIS AND USES THEREOF
; FILE REFERENCE: CLFR:244US
; CURRENT APPLICATION NUMBER: US/11/066,648A
; CURRENT FILING DATE: 2005-02-25
; PRIOR APPLICATION NUMBER: 60/547,499
; PRIOR FILING DATE: 2004-02-25
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1044
; TYPE: DNA
; ORGANISM: Ehrlichia canis
US-11-066-648A-1

```

Query Match	5.6%	Score 45.2;	DB 7;	Length 1044;
Best Local Similarity	47.8%;	Pred. No. 0.19;		
Matches 131;	Conservative 0;	Mismatches 143;	Indels 0;	Gaps 0;
Qy	82	AAATTTTCTGAGGGAGGAGGAC	TTTATATAGGTAGTCAGTATAAAGTTGGTATTCC	141
Db	679	AGTTATTATT	TTGGTAGAATTTGCAAGCTTCTGATAAGAAAGTGATCAAGCTGTAGTAAAA	738
Qy	142	AAATTTTAGTAAATTTTTCAGCTGGAAGAAACAATTCCTCGTATTACAAAAAGACATTTTTCG	201	
Db	739	AAGCTTGGTATTTTTTTTCCCTTAATCAAGAGACTACTCGTACTATGATCAACATTAGTGGT	798	
Qy	202	TTAGGTCTTGATAGTCTGAGATAAAATACCTCACAGCAATTTTACACGATCATATGACCC	261	
Db	799	GGTGCTGTTACAAAGCATGCAAAAATTAACAAATGCTATAAAATTTATGGAAATTCCTG	858	
Qy	262	ACTTATGCAAGCAGTTTTTGCAGGGTTTTAGTGGTATCATTTGGATATTATGTTAATGACTTTT	321	
Db	859	ACTAGTGTAAAGGACACAGAAAGTTTATGCTCAAGTTAATCAAGAATACCCCTATTGTGAA	918	
Qy	322	AGGTAGAAATTTGAAGTCTCTATGAGAAATTTG	355	
Db	919	GGTGAGAACTTTTCAGAAAGTATTAAAGACTTTTG	952	

```

RESULT 7
US-10-750-185-58930/c
; Sequence 58930, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58930
; LENGTH: 1773
; TYPE: DNA
; ORGANISM: Bovine 19866880962402
US-10-750-185-58930

```



```
; LENGTH: 31028
; TYPE: DNA
; ORGANISM: Coronavirus
US-10-829-826B-22

Query Match      5.1%; Score 41.2; DB 6; Length 31028;
Best Local Similarity 51.1%; Pred. No. 4.3; Mismatches 0; Indels 0; Gaps 0;
Matches 97; Conservative 0;

Qy 291 TGGTATCATTTGGATATTATGTTAATGACTTTTAGGGTAGAATTTGAAGGTTCTTTATGAGAA 350
Db 10449 TAGTACTGGTTGTGCATACCTGCTACTGATTTCAATGGGGATTTTATGTCCTTTAAGGA 10508

Qy 351 TTTTGAACCTGAAAGACAATGGTACCTGAGAAATAGCCCAAGCTACAAATTTTTCGCTTT 410
Db 10509 TGCTCAGGTTGTCCAAATGGCCGTTTCAGGATTATACAACTCTGTTAAATTTTGTAGCATG 10568

Qy 411 GTCTCGAAATGCTACAAATAGTGTAAATAGTTTATAGTACTAGAGAAATACCGCGTTGT 470
Db 10569 GCTTTATGCTGCTATACCTTAATAATGTAATTTGGTTTGTACAAAGTGATAAGTGTCTGT 10628

Qy 471 TGACAAGTCT 480
Db 10629 TGAAGATTTT 10638

RESULT 11
US-10-829-826B-26
; Sequence 26, Application US/10829826B
; Publication No. US20050266397A1
; GENERAL INFORMATION:
; APPLICANT: Ecker, David J.
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Blyn, Lawrence B.
; APPLICANT: Hall, Thomas A.
; APPLICANT: Massire, Christian
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES
; FILE REFERENCE: IBIS0075-100 (DIBIS-0058US)
; CURRENT APPLICATION NUMBER: US/10/829,826B
; CURRENT FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 31028
; TYPE: DNA
; ORGANISM: Coronavirus
US-10-829-826B-26

Query Match      5.1%; Score 41.2; DB 6; Length 31028;
Best Local Similarity 51.1%; Pred. No. 4.3; Mismatches 0; Indels 0; Gaps 0;
Matches 97; Conservative 0;

Qy 291 TGGTATCATTTGGATATTATGTTAATGACTTTTAGGGTAGAATTTGAAGGTTCTTTATGAGAA 350
Db 10449 TAGTACTGGTTGTGCATACCTGCTACTGATTTCAATGGGGATTTTATGTCCTTTAAGGA 10508

Qy 351 TTTTGAACCTGAAAGACAATGGTACCTGAGAAATAGCCCAAGCTACAAATTTTTCGCTTT 410
Db 10509 TGCTCAGGTTGTCCAAATGGCCGTTTCAGGATTATACAACTCTGTTAAATTTTGTAGCATG 10568

Qy 411 GTCTCGAAATGCTACAAATAGTGTAAATAGTTTATAGTACTAGAGAAATACCGCGTTGT 470
Db 10569 GCTTTATGCTGCTATACCTTAATAATGTAATTTGGTTTGTACAAAGTGATAAGTGTCTGT 10628

Qy 471 TGACAAGTCT 480
Db 10629 TGAAGATTTT 10638

RESULT 12
US-10-829-826B-23
; Sequence 23, Application US/10829826B
; Publication No. US20050266397A1
; GENERAL INFORMATION:
; APPLICANT: Ecker, David J.
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Blyn, Lawrence B.
; APPLICANT: Hall, Thomas A.
; APPLICANT: Massire, Christian
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES
; FILE REFERENCE: IBIS0075-100 (DIBIS-0058US)
; CURRENT APPLICATION NUMBER: US/10/829,826B
; CURRENT FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 31032
; TYPE: DNA
; ORGANISM: Coronavirus
US-10-829-826B-23

Query Match      5.1%; Score 41.2; DB 6; Length 31032;
Best Local Similarity 51.1%; Pred. No. 4.3; Mismatches 0; Indels 0; Gaps 0;
Matches 97; Conservative 0;

Qy 291 TGGTATCATTTGGATATTATGTTAATGACTTTTAGGGTAGAATTTGAAGGTTCTTTATGAGAA 350
Db 10449 TAGTACTGGTTGTGCATACCTGCTACTGATTTCAATGGGGATTTTATGTCCTTTAAGGA 10508

Qy 351 TTTTGAACCTGAAAGACAATGGTACCTGAGAAATAGCCCAAGCTACAAATTTTTCGCTTT 410
Db 10509 TGCTCAGGTTGTCCAAATGGCCGTTTCAGGATTATACAACTCTGTTAAATTTTGTAGCATG 10568

Qy 411 GTCTCGAAATGCTACAAATAGTGTAAATAGTTTATAGTACTAGAGAAATACCGCGTTGT 470
Db 10569 GCTTTATGCTGCTATACCTTAATAATGTAATTTGGTTTGTACAAAGTGATAAGTGTCTGT 10628

Qy 471 TGACAAGTCT 480
Db 10629 TGAAGATTTT 10638

RESULT 13
US-10-829-826B-24
; Sequence 24, Application US/10829826B
; Publication No. US20050266397A1
; GENERAL INFORMATION:
; APPLICANT: Ecker, David J.
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Blyn, Lawrence B.
; APPLICANT: Hall, Thomas A.
; APPLICANT: Massire, Christian
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES
; FILE REFERENCE: IBIS0075-100 (DIBIS-0058US)
; CURRENT APPLICATION NUMBER: US/10/829,826B
; CURRENT FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 31100
; TYPE: DNA
; ORGANISM: Coronavirus
US-10-829-826B-24

Query Match      5.1%; Score 41.2; DB 6; Length 31100;
Best Local Similarity 51.1%; Pred. No. 4.3; Mismatches 0; Indels 0; Gaps 0;
Matches 97; Conservative 0;

Qy 291 TGGTATCATTTGGATATTATGTTAATGACTTTTAGGGTAGAATTTGAAGGTTCTTTATGAGAA 350
Db 10449 TAGTACTGGTTGTGCATACCTGCTACTGATTTCAATGGGGATTTTATGTCCTTTAAGGA 10508

Qy 351 TTTTGAACCTGAAAGACAATGGTACCTGAGAAATAGCCCAAGCTACAAATTTTTCGCTTT 410
```

Db 10509 TGCTCAGGTGTGCAATTTGCCGCTTCAGGATTATATACAAATCTGTTAAATTTTGTAGCATG 10568  
Qy 411 GTCTCGAATGCTACAATAGTATGAATTAAGTTTATAGTACTAGAGAAATAACGGCGTTGT 470  
Db 10569 GCTTTATGCTGCTACTACTTAATAATTTGTAATGGTTTGTACAAAGTGATAAGTGTCTGT 10628  
Qy 471 TGACAAAGTCT 480  
Db 10629 TGAAGATTTT 10638

RESULT 14  
US-10-240-708-11  
; Sequence 11, Application US/10240708  
; Publication No. US20050282157A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication  
; TITLE OF INVENTION: by Assessing DNA Methylation  
; FILE REFERENCE: 5013.1012  
; CURRENT APPLICATION NUMBER: US/10/240,708  
; PRIOR FILING DATE: 2002-10-03  
; PRIOR APPLICATION NUMBER: PCT/EP01/03971  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: DE 10019058.8  
; PRIOR FILING DATE: 2000-04-06  
; PRIOR APPLICATION NUMBER: DE 10019173.8  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 98  
; SEQ ID NO 11  
; LENGTH: 6317  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-240-708-11

Query Match 4.9%; Score 40; DB 6; Length 6317;  
Best Local Similarity 44.8%; Pred. No. 4.9; Mismatches 190; Indels 0; Gaps 0;  
Matches 154; Conservative 0;  
Qy 12 AAGATTGTTGTAGGTGTACGCTGAGTACATTTGTTTTTCTTATCTGATGGTCTTT 71  
Db 1588 AAGTTTTTTTATTTTATTTTATGAATATTTTAAATATATAGAGAAATGGAAGAGTAT 1647  
Qy 72 TCTCGATGCAAAATTTTCTGAAGGAGGAGGAGACTTTATATATAGGTAGTCAGTATAAAGT 131  
Db 1648 TGTATAAATAATATAATGAATAGATTTTAATAATTTGTTAAATTTTGTCTGATGTATT 1707  
Qy 132 TGGTATTCCTCAATTTTATGTAATTTTTCAGCTGAAGAACAAATTCCTGGTATTACAAAAA 191  
Db 1708 TAATATTTTAAATTTTATGATTAATTAATTTTAAAGTATTTTATAAATATTATAAGATA 1767  
Qy 192 GATTTTTGCGTTAGGTCTTGATAAGTCTGAGATAAAATATCTCACAGCAATTTTACACGATC 251  
Db 1768 TTTTATTTTAAATATTTTATGATATGATATTTTAAAGTAAGGATAAATTTATATAGATAT 1827  
Qy 252 ATATGACCCCTACTTATGCAAGAGAGTTTTCAGGGTTTATGGGTATCATTGGATATTATGT 311  
Db 1828 AATTATATTATAATTTTAAAGAAAAATTTTATAATTTTAAATATTATTTTCAAAAATTAAT 1887  
Qy 312 TAATGACTTTTAGGGTAGAATTTTCAGGCTTCTTATGAGAAATTTTG 355  
Db 1888 TTTTATTTTATTTATTTGTTTAAATGTTTATAGTTATTTTG 1931

RESULT 15  
US-10-750-185-39647  
; Sequence 39647, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIN version 3.1  
; SEQ ID NO 39647  
; LENGTH: 1289  
; TYPE: DNA  
; ORGANISM: Bovine 19866880742947  
US-10-750-185-39647  
Query Match 4.8%; Score 39.2; DB 6; Length 1289;  
Best Local Similarity 48.6%; Pred. No. 4.6; Mismatches 113; Indels 0; Gaps 0;  
Matches 107; Conservative 0;  
Qy 295 ATCATTGGATATTATGTTAATGACTTTTAGGGTAGAATTTGAAGGTTCTTTATGAGAAATTT 354  
Db 193 ATTATTTTGGTTTATTATTAACAGGATGAATTAACCTTGGCTGATTTCTGAAGTGGATAAC 252  
Qy 355 GAACCTGAAAGACAATGGTACCCCTGAGAAATAGCCAAAGCTACAAATTTTTTTCCTTTGCT 414  
Db 253 CAAAAGCGAGGGAACCGCATTTATGAAGACAAGCAAGAAAGACACTTGGATACTTTAAAT 312  
Qy 415 CGAATGCTACAAATAGTGAATAAGTTTATAGTACTAGAGAAATAACGGCTTCTTGTAC 474  
Db 313 AAAAAAGAACGAGAACTAGATATGAAGAAAAAGAGCTAGAGGTTTATCCATGATTTTATA 372  
Qy 475 AAGTCTCTTAATGTTAAATGTTTGTATTGATATTGCTAGTG 514  
Db 373 AAACCTTTTATTACACCTTAGTCATGATTTATTAAGTG 412  
Search completed: December 28, 2005, 15:34:25  
Job time : 842 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 28, 2005, 12:12:19 ; Search time 868 Seconds  
(without alignments)  
7745.397 Million cell updates/sec

Title: US-10-731-554-45  
Perfect score: 813  
Sequence: 1 atgaattacaaagattgt.....ttggattgattgctgcta 813

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues  
Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA Main:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	813	100.0	813	3	US-09-811-007-45
2	813	100.0	813	5	US-10-062-624-45
3	813	100.0	813	5	US-10-062-051-45
4	813	100.0	813	5	US-10-062-920-45
5	813	100.0	813	5	US-10-680-349-45
6	813	100.0	813	8	US-10-731-554-45
7	162.2	20.0	840	5	US-10-059-964-23
8	162.2	20.0	840	5	US-10-314-639-23
9	162.2	20.0	840	8	US-10-901-714-23
10	162.2	20.0	840	8	US-10-901-714-23
11	150.8	18.5	852	5	US-10-059-964-25
12	150.8	18.5	852	5	US-10-314-639-25
13	150.8	18.5	852	8	US-10-901-714-25
14	150.8	18.5	852	8	US-10-901-714-25
15	147.2	18.1	840	3	US-09-811-007-41
16	147.2	18.1	840	5	US-10-062-624-41
17	147.2	18.1	840	5	US-10-062-051-41
18	147.2	18.1	840	5	US-10-062-920-41
19	147.2	18.1	840	8	US-10-680-349-41
20	147.2	18.1	840	8	US-10-731-554-41
21	147.2	18.1	843	5	US-10-059-964-47
22	147.2	18.1	843	5	US-10-314-639-47
23	147.2	18.1	843	8	US-10-901-714-47

24	147.2	18.1	843	8	US-10-901-774-47	Sequence 47, Appl
25	146.6	18.0	726	5	US-10-059-964-61	Sequence 61, Appl
26	146.6	18.0	726	5	US-10-314-639-61	Sequence 61, Appl
27	146.6	18.0	726	5	US-10-901-714-61	Sequence 61, Appl
28	146.6	18.0	726	5	US-10-901-774-61	Sequence 61, Appl
29	146.6	18.0	849	9	US-10-138-163-48	Sequence 48, Appl
30	135	16.6	828	5	US-10-059-964-27	Sequence 27, Appl
31	135	16.6	828	5	US-10-314-639-27	Sequence 27, Appl
32	135	16.6	828	8	US-10-901-714-27	Sequence 27, Appl
33	135	16.6	828	8	US-10-901-774-27	Sequence 27, Appl
34	133.8	16.5	840	5	US-10-059-964-59	Sequence 59, Appl
35	133.8	16.5	840	5	US-10-314-639-59	Sequence 59, Appl
36	133.8	16.5	840	8	US-10-901-714-59	Sequence 59, Appl
37	133.8	16.5	840	8	US-10-901-774-59	Sequence 59, Appl
38	129.6	15.9	843	5	US-10-059-964-5	Sequence 5, Appl
39	129.6	15.9	843	5	US-10-314-639-5	Sequence 5, Appl
40	129.6	15.9	843	8	US-10-901-714-5	Sequence 5, Appl
41	129.6	15.9	843	8	US-10-901-774-5	Sequence 5, Appl
42	126.4	15.5	837	5	US-10-062-994-9	Sequence 9, Appl
43	126.4	15.5	837	5	US-10-059-964-9	Sequence 9, Appl
44	126.4	15.5	837	5	US-10-062-994-9	Sequence 9, Appl
45	126.4	15.5	837	5	US-10-314-639-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1

US-09-811-007-45  
; Sequence 45, Application US/09811007  
; Publication No. US20030185849A1  
; GENERAL INFORMATION:  
; APPLICANT: Walker, David H.  
; APPLICANT: McBride, Jere W.  
; APPLICANT: Yu, Xue-Jie  
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein  
; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof  
; FILE REFERENCE: D6152C1F2  
; CURRENT APPLICATION NUMBER: US/09/811,007  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 09/660,587  
; PRIOR FILING DATE: 2000-09-12  
; NUMBER OF SEQ ID NOS: 46  
; SEQ ID NO 45  
; LENGTH: 813  
; TYPE: DNA  
; ORGANISM: Ehrlichia canis  
; FEATURE:  
; OTHER INFORMATION: nucleic acid sequence of E. canis p28-9  
US-09-811-007-45

Query Match	100.0%	Score 813;	DB 3;	Length 813;
Best Local Similarity	100.0%	Pred. No. 9,2e-181;		
Matches 813;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGAATTACAAAAGATTGTTGTTAGGTGTTACGCTGAGTACATTTGTTTTTCTTATCT	60	
Db	1	ATGAATTACAAAAGATTGTTGTTAGGTGTTACGCTGAGTACATTTGTTTTTCTTATCT	60	
Qy	61	GATCGTCTTTTCTGATGCAATTTTCTGAAGGGAGGAGGACCTTTATATAGGTAGT	120	
Db	61	GATCGTCTTTTCTGATGCAATTTTCTGAAGGGAGGAGGACCTTTATATAGGTAGT	120	
Qy	121	CAGTATAAAGTGTGTTATCCCAATTTTGTAGTAAATTTTTCAGCTGAAGAAACAATTCCTGGT	180	
Db	121	CAGTATAAAGTGTGTTATCCCAATTTTGTAGTAAATTTTTCAGCTGAAGAAACAATTCCTGGT	180	
Qy	181	ATTACAAAAAGATTTTTCGTTAGGTCTTTGATAAGTCTGAGATAAATCTACAGCAAT	240	
Db	181	ATTACAAAAAGATTTTTCGTTAGGTCTTTGATAAGTCTGAGATAAATCTACAGCAAT	240	
Qy	241	TTTACAGATCATATGACCTTACTTATGCAAGCAATTTTCAGGTTTGTAGGTATCATTT	300	

```
Db      241  TTTACACGATCATATGACCCCTACTTATGCAAGCAGTTTTTGCAGGGTTTAGTGGTATCATTT 300
Qy      301  GGTATATTGTTAATGACCTTTAGGTTAGATTTGAGGTTCTTATGAGATTTTGAACCT 360
Db      301  GGTATATTGTTAATGACCTTTAGGTTAGATTTGAGGTTCTTATGAGATTTTGAACCT 360
Qy      361  GAAAGACAATGGTACCTCGAGATAGCCAAAGCTACAAAATTTTTTCTTTGTTCTCGAAAT 420
Db      361  GAAAGACAATGGTACCTCGAGATAGCCAAAGCTACAAAATTTTTTCTTTGTTCTCGAAAT 420
Qy      421  GCTACAAAATAGTGATTAATAGTTTATAGTACTAGAGAAATAACGGCGTTGTTGACAAAGTCT 480
Db      421  GCTACAAAATAGTGATTAATAGTTTATAGTACTAGAGAAATAACGGCGTTGTTGACAAAGTCT 480
Qy      481  CTTAAATGTAATGTTTGTATGATATTGCTAGTGGTAGTATTCCTTTAGCACCTTATATG 540
Db      481  CTTAAATGTAATGTTTGTATGATATTGCTAGTGGTAGTATTCCTTTAGCACCTTATATG 540
Qy      541  TGTGCTGGTGTTCGTCAGATTTATATAAAGTTTTTTAGGTATATCATTTGCTTAAGTTTTCT 600
Db      541  TGTGCTGGTGTTCGTCAGATTTATATAAAGTTTTTTAGGTATATCATTTGCTTAAGTTTTCT 600
Qy      601  TATCAAGTTAAAGTTTGGTGTCAACTACCTCTAAATGTTAAATGTTAAATGTTAAATGTTAA 660
Db      601  TATCAAGTTAAAGTTTGGTGTCAACTACCTCTAAATGTTAAATGTTAAATGTTAAATGTTAA 660
Qy      661  GGTATTTACCATAGGTTGAGTGATAGGCATGAGAGAGTAGAAGATAGCTTACCACCTCT 720
Db      661  GGTATTTACCATAGGTTGAGTGATAGGCATGAGAGAGTAGAAGATAGCTTACCACCTCT 720
Qy      721  ACTGCATTATCTGACGTTCTCAGAACTACTTCAAGCTTCTGCTACTTTTAAATACGATTTAT 780
Db      721  ACTGCATTATCTGACGTTCTCAGAACTACTTCAAGCTTCTGCTACTTTTAAATACGATTTAT 780
Qy      781  TTTGGTTGGAGATTGGATTTAGATTTGGCGCTA 813
Db      781  TTTGGTTGGAGATTGGATTTAGATTTGGCGCTA 813
```

## RESULT 2

```
US-10-062-624-45
; Sequence 45, Application US/10062624
; Publication No. US20020115840A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-Kilodalton Immunodominant Protein
; FILE REFERENCE: D6152CIP2/D1
; CURRENT APPLICATION NUMBER: US/10/062,624
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 09/660,587
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 45
; LENGTH: 813
; TYPE: DNA
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: nucleic acid sequence of E. canis p28-9
US-10-062-624-45
```

```
Query Match      100.0%; Score 813; DB 5; Length 813;
Best Local Similarity 100.0%; Pred. No. 9.2e-181;
Matches 813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  ATGAATTTACAAAAGATTTGTTGAGGTGTTTACCTCAGTACATTTCTTTTCTTATCT 60
Db      1  ATGAATTTACAAAAGATTTGTTGAGGTGTTTACCTCAGTACATTTCTTTTCTTATCT 60
Qy      61  GATGGTCTTTTCTCGATGCAAAATTTTCTGAAAGGAGGAGGACTTTATATAGTACT 120
```

```
Db      61  GATGGTCTTTTCTCGATGCAAAATTTTCTGAAAGGAGGAGGACTTTATATAGTACT 120
Qy      121  CAGTATTAAGTTGGTATTTCCCAATTTTCTAGTAAATTTTTCAGCTGCAAGAACAAATTTCTGGT 180
Db      121  CAGTATTAAGTTGGTATTTCCCAATTTTCTAGTAAATTTTTCAGCTGCAAGAACAAATTTCTGGT 180
Qy      181  ATTACAAAAGATTTTTCGCGTTAGGTTCTTGATAAGTCTGAGATAAAATCTCACAGCAAT 240
Db      181  ATTACAAAAGATTTTTCGCGTTAGGTTCTTGATAAGTCTGAGATAAAATCTCACAGCAAT 240
Qy      241  TTTACACGATCATATGACCCCTACTTATGCAAGCAGCTTTTTCAGGCTTTAGTGGTATCATTT 300
Db      241  TTTACACGATCATATGACCCCTACTTATGCAAGCAGCTTTTTCAGGCTTTAGTGGTATCATTT 300
Qy      301  GGATATTTATGTTAAATGACTTTTAGGGTAGAATTTTGAAGGTTCTTATGAGAAATTTTCAACCT 360
Db      301  GGATATTTATGTTAAATGACTTTTAGGGTAGAATTTTGAAGGTTCTTATGAGAAATTTTCAACCT 360
Qy      361  GAAAGACAATGGTACCCCTGAGAAATAGCCAAAGCTACAAAATTTTTTGTCTCGAAAT 420
Db      361  GAAAGACAATGGTACCCCTGAGAAATAGCCAAAGCTACAAAATTTTTTGTCTCGAAAT 420
Qy      421  GCTACAAAATAGTGATTAATAGTTTATAGTACTAGAGAAATAACGGCGTTGTTGACAAAGTCT 480
Db      421  GCTACAAAATAGTGATTAATAGTTTATAGTACTAGAGAAATAACGGCGTTGTTGACAAAGTCT 480
Qy      481  CTTAAATGTAATGTTTGTATGATATTGCTAGTGGTAGTATTCCTTTAGCACCTTATATG 540
Db      481  CTTAAATGTAATGTTTGTATGATATTGCTAGTGGTAGTATTCCTTTAGCACCTTATATG 540
Qy      541  TGTGCTGGTGTTCGTCAGATTTATATAAAGTTTTTTAGGTATATCATTTGCTTAAGTTTTCT 600
Db      541  TGTGCTGGTGTTCGTCAGATTTATATAAAGTTTTTTAGGTATATCATTTGCTTAAGTTTTCT 600
Qy      601  TATCAAGTTAAAGTTTGGTGTCAACTACCTCTAAATGTTAAATGTTAAATGTTAAATGTTAA 660
Db      601  TATCAAGTTAAAGTTTGGTGTCAACTACCTCTAAATGTTAAATGTTAAATGTTAAATGTTAA 660
Qy      661  GGTATTTACCATAGGTTGAGTGATAGGCATGAGAGAGTAGAAGATAGCTTACCACCTCT 720
Db      661  GGTATTTACCATAGGTTGAGTGATAGGCATGAGAGAGTAGAAGATAGCTTACCACCTCT 720
Qy      721  ACTGCATTATCTGACGTTCTCAGAACTACTTCAAGCTTCTGCTACTTTTAAATACGATTTAT 780
Db      721  ACTGCATTATCTGACGTTCTCAGAACTACTTCAAGCTTCTGCTACTTTTAAATACGATTTAT 780
Qy      781  TTTGGTTGGAGATTGGATTTAGATTTGGCGCTA 813
Db      781  TTTGGTTGGAGATTGGATTTAGATTTGGCGCTA 813
```

## RESULT 3

```
US-10-062-051-45
; Sequence 45, Application US/10062051
; Publication No. US20030073095A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
; FILE REFERENCE: D6152CIP2
; CURRENT APPLICATION NUMBER: US/10/062,051
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US/09/660,587
; PRIOR FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 09/261,358
; PRIOR FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 45
; LENGTH: 813
; TYPE: DNA
; ORGANISM: Ehrlichia canis
```

```
; FEATURE:
; OTHER INFORMATION: nucleic acid sequence of E. canis p28-9
US-10-062-051-45

Query Match      100.0%; Score 813; DB 5; Length 813;
Best Local Similarity 100.0%; Pred. No. 9.2e-181;
Matches 813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAATTACAAAAGATTGTTGTAGGTGTTACGCTGAGTACATTTGTTTTTCTTATCT 60
Db 1 ATGAATTACAAAAGATTGTTGTAGGTGTTACGCTGAGTACATTTGTTTTTCTTATCT 60

Qy 61 GATGTCCTTTTCTGATGCAAAATTTTCTGAAGGGAGGAGGACTTTATATAGGTAGT 120
Db 61 GATGTCCTTTTCTGATGCAAAATTTTCTGAAGGGAGGAGGACTTTATATAGGTAGT 120

Qy 121 CAGTATAAAGTTGGTATCCCAATTTTAGTAAATTTTTCAGCTGAAGAAACAAATTCCTGGT 180
Db 121 CAGTATAAAGTTGGTATCCCAATTTTAGTAAATTTTTCAGCTGAAGAAACAAATTCCTGGT 180

Qy 181 ATTAACAAAAGATTGTTGCGTTAGGTCTTGATAGTCTGAGATAAATCTCACAGCAAT 240
Db 181 ATTAACAAAAGATTGTTGCGTTAGGTCTTGATAGTCTGAGATAAATCTCACAGCAAT 240

Qy 241 TTTACAGATCATATGACCCCTACTTATGCAAGAGTTTTCGAGGTTTAGTGGTATCAT 300
Db 241 TTTACAGATCATATGACCCCTACTTATGCAAGAGTTTTCGAGGTTTAGTGGTATCAT 300

Qy 301 GGATATTATGTTAATGACTTTAGGTAGAAATTTCAAGGTTCTTATGAGAAATTTTGAACCT 360
Db 301 GGATATTATGTTAATGACTTTAGGTAGAAATTTCAAGGTTCTTATGAGAAATTTTGAACCT 360

Qy 361 GAAAGACAAATGCTGAGAAATAGCCAAAGCTACAAAATTTTTCGCTGCTCGAAAT 420
Db 361 GAAAGACAAATGCTGAGAAATAGCCAAAGCTACAAAATTTTTCGCTGCTCGAAAT 420

Qy 421 GCTACAAATAGTGAATTAAGTTTATAGTACTAGAGAAATACCGCGTTGTTGCAAGTCT 480
Db 421 GCTACAAATAGTGAATTAAGTTTATAGTACTAGAGAAATACCGCGTTGTTGCAAGTCT 480

Qy 481 CTTAATGTAATGTTTATGATATGCTAGTGGTAGTATTCCTTTAGCACCCTTATATG 540
Db 481 CTTAATGTAATGTTTATGATATGCTAGTGGTAGTATTCCTTTAGCACCCTTATATG 540

Qy 541 TGTGCTGGTCTGGTGCAGATTATATAAAGTTTTTAGGTATATCATTTGCCCTAAGTTTCT 600
Db 541 TGTGCTGGTCTGGTGCAGATTATATAAAGTTTTTAGGTATATCATTTGCCCTAAGTTTCT 600

Qy 601 TATCAAGTTAAGTTTGGTGTCAACTACCCCTCTAAATGTTAATCTATGTTTGGTGGG 660
Db 601 TATCAAGTTAAGTTTGGTGTCAACTACCCCTCTAAATGTTAATCTATGTTTGGTGGG 660

Qy 661 GGTATTACCAATAGTTGTTAGGTGATAGGATGAGAGATGAGAAATAGCTTACCCTCT 720
Db 661 GGTATTACCAATAGTTGTTAGGTGATAGGATGAGAGATGAGAAATAGCTTACCCTCT 720

Qy 721 ACTGCAATATCTGACGTTCTAGAACTACTTTCAGCTTCTGCTACTCTTAAATACTGATTAT 780
Db 721 ACTGCAATATCTGACGTTCTAGAACTACTTTCAGCTTCTGCTACTCTTAAATACTGATTAT 780

Qy 781 TTTGGTTGGGAGTTGGATTAGATTTCGGCTA 813
Db 781 TTTGGTTGGGAGTTGGATTAGATTTCGGCTA 813
```

## RESULT 4

```
US-10-062-920-45
; Sequence 45, Application US/10062920
; Publication No. US20030096250A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBrine, Jere W.
; APPLICANT: Yu, Xue-Jie
```

```
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
; FILE REFERENCE: D6152CIP2
; CURRENT APPLICATION NUMBER: US/10/062,920
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US/09/660,587
; PRIOR FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 09/261,358
; PRIOR FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 45
; LENGTH: 813
; TYPE: DNA
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: nucleic acid sequence of E. canis p28-9
US-10-062-920-45
```

```
Query Match      100.0%; Score 813; DB 5; Length 813;
Best Local Similarity 100.0%; Pred. No. 9.2e-181;
Matches 813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAATTACAAAAGATTGTTGTAGGTGTTACGCTGAGTACATTTGTTTTTCTTATCT 60
Db 1 ATGAATTACAAAAGATTGTTGTAGGTGTTACGCTGAGTACATTTGTTTTTCTTATCT 60

Qy 61 GATGTCCTTTTCTGATGCAAAATTTTCTGAAGGGAGGAGGACTTTATATAGGTAGT 120
Db 61 GATGTCCTTTTCTGATGCAAAATTTTCTGAAGGGAGGAGGACTTTATATAGGTAGT 120

Qy 121 CAGTATAAAGTTGGTATCCCAATTTTAGTAAATTTTTCAGCTGAAGAAACAAATTCCTGGT 180
Db 121 CAGTATAAAGTTGGTATCCCAATTTTAGTAAATTTTTCAGCTGAAGAAACAAATTCCTGGT 180

Qy 181 ATTAACAAAAGATTGTTGCGTTAGGTCTTGATAGTCTGAGATAAATCTCACAGCAAT 240
Db 181 ATTAACAAAAGATTGTTGCGTTAGGTCTTGATAGTCTGAGATAAATCTCACAGCAAT 240

Qy 241 TTTACAGATCATATGACCCCTACTTATGCAAGAGTTTTCGAGGTTTAGTGGTATCAT 300
Db 241 TTTACAGATCATATGACCCCTACTTATGCAAGAGTTTTCGAGGTTTAGTGGTATCAT 300

Qy 301 GGATATTATGTTAATGACTTTAGGTAGAAATTTCAAGGTTCTTATGAGAAATTTTGAACCT 360
Db 301 GGATATTATGTTAATGACTTTAGGTAGAAATTTCAAGGTTCTTATGAGAAATTTTGAACCT 360

Qy 361 GAAAGACAAATGCTGAGAAATAGCCAAAGCTACAAAATTTTTCGCTGCTCGAAAT 420
Db 361 GAAAGACAAATGCTGAGAAATAGCCAAAGCTACAAAATTTTTCGCTGCTCGAAAT 420

Qy 421 GCTACAAATAGTGAATTAAGTTTATAGTACTAGAGAAATACCGCGTTGTTGCAAGTCT 480
Db 421 GCTACAAATAGTGAATTAAGTTTATAGTACTAGAGAAATACCGCGTTGTTGCAAGTCT 480

Qy 481 CTTAATGTAATGTTTATGATATGCTAGTGGTAGTATTCCTTTAGCACCCTTATATG 540
Db 481 CTTAATGTAATGTTTATGATATGCTAGTGGTAGTATTCCTTTAGCACCCTTATATG 540

Qy 541 TGTGCTGGTCTGGTGCAGATTATATAAAGTTTTTAGGTATATCATTTGCCCTAAGTTTCT 600
Db 541 TGTGCTGGTCTGGTGCAGATTATATAAAGTTTTTAGGTATATCATTTGCCCTAAGTTTCT 600

Qy 601 TATCAAGTTAAGTTTGGTGTCAACTACCCCTCTAAATGTTAATCTATGTTTGGTGGG 660
Db 601 TATCAAGTTAAGTTTGGTGTCAACTACCCCTCTAAATGTTAATCTATGTTTGGTGGG 660

Qy 661 GGTATTACCAATAGTTGTTAGGTGATAGGATGAGAGATGAGAAATAGCTTACCCTCT 720
Db 661 GGTATTACCAATAGTTGTTAGGTGATAGGATGAGAGATGAGAAATAGCTTACCCTCT 720

Qy 721 ACTGCAATATCTGACGTTCTAGAACTACTTTCAGCTTCTGCTACTCTTAAATACTGATTAT 780
Db 721 ACTGCAATATCTGACGTTCTAGAACTACTTTCAGCTTCTGCTACTCTTAAATACTGATTAT 780

Qy 781 TTTGGTTGGGAGTTGGATTAGATTTCGGCTA 813
Db 781 TTTGGTTGGGAGTTGGATTAGATTTCGGCTA 813
```

```
QY 781 TTTGGTTGGAGATTGGATTAGATTGGCGTA 813
Db |||||
781 TTTGGTTGGAGATTGGATTAGATTGGCGTA 813
|||

RESULT 5
US-10-680-349-45
; Sequence 45, Application US/10680349
; Publication No. US20040198951A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-Kilodalton Immunodominant Protein
; FILE REFERENCE: D6152CIP2/D1
; CURRENT APPLICATION NUMBER: US/10/680,349
; CURRENT FILING DATE: 2003-10-07
; PRIOR APPLICATION NUMBER: US/10/062,624
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 09/660,587
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 45
; LENGTH: 813
; TYPE: DNA
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: nucleic acid sequence of E. canis p28-9
US-10-680-349-45

Query Match 100.0%; Score 813; DB 8; Length 813;
Best Local Similarity 100.0%; Pred. No. 9.2e-181;
Matches 813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAATTTACAAAAGATTGTTGAGGTGTTACCTCAGTACATTTGTTTTTCTTATCT 60
Db |||||
1 ATGAATTTACAAAAGATTGTTGAGGTGTTACCTCAGTACATTTGTTTTTCTTATCT 60
|||

QY 61 GATGGTGTCTTTTCTGATGCAAAATTTTCTGAAGGAGGAGGAGGACTTTATAGTAGT 120
Db |||||
61 GATGGTGTCTTTTCTGATGCAAAATTTTCTGAAGGAGGAGGAGGACTTTATAGTAGT 120
|||

QY 121 CAGTATAAAGTTGGTATCCCAATTTTATAGTAATTTTTCAGCTGAAGAAACAATTCCTGT 180
Db |||||
121 CAGTATAAAGTTGGTATCCCAATTTTATAGTAATTTTTCAGCTGAAGAAACAATTCCTGT 180
|||

QY 121 CAGTATAAAGTTGGTATCCCAATTTTATAGTAATTTTTCAGCTGAAGAAACAATTCCTGT 180
Db |||||
121 CAGTATAAAGTTGGTATCCCAATTTTATAGTAATTTTTCAGCTGAAGAAACAATTCCTGT 180
|||

QY 181 ATTACAAAAGATTGTTGCGTTAGTCTTGATTAAGTCTTGAGATAAATACTCACAGCAAT 240
Db |||||
181 ATTACAAAAGATTGTTGCGTTAGTCTTGATTAAGTCTTGAGATAAATACTCACAGCAAT 240
|||

QY 241 TTTACACGATCATATGACCCCTACTTATGCAAGCACTTTTGCAGGGTTAGTGTATCAT 300
Db |||||
241 TTTACACGATCATATGACCCCTACTTATGCAAGCACTTTTGCAGGGTTAGTGTATCAT 300
|||

QY 301 GGATATTATGTTAAATGACTTTTATAGGTTAGGTTAGGTTAGGTTAGGTTAGGTTAGG 360
Db |||||
301 GGATATTATGTTAAATGACTTTTATAGGTTAGGTTAGGTTAGGTTAGGTTAGGTTAGG 360
|||

QY 361 GAAAGACAAATGGTACCTCAGATAGCCAAAGCTACAAATTTTTCGTTTGTCTCGAAAT 420
Db |||||
361 GAAAGACAAATGGTACCTCAGATAGCCAAAGCTACAAATTTTTCGTTTGTCTCGAAAT 420
|||

QY 421 GCTACAAATAGTGATTAATAGTTTATAGTACTAGAGAAATAACGGCGTTGTTGACAGTCT 480
Db |||||
421 GCTACAAATAGTGATTAATAGTTTATAGTACTAGAGAAATAACGGCGTTGTTGACAGTCT 480
|||

QY 481 CTTAAATGTAATGTTGTTATGATATTTGCTAGTGGTAGTATTCCTTTAGCAGCTTATAG 540
Db |||||
481 CTTAAATGTAATGTTGTTATGATATTTGCTAGTGGTAGTATTCCTTTAGCAGCTTATAG 540
|||

QY 541 TGTGCTGGTGTGGTGAGATTATATAAAGTTTTTATAGGTATATCATTTGCGCTAAGTTTCT 600
|||
```

```
Db 541 TGTGCTGGTGTGGTGAGATTATATAAAGTTTTTATAGGTATATCATTTGCGCTAAGTTTCT 600
QY 601 TATCAAGTTAAGTTTGGTGTCACTACCTCTAAATGTTAATACTATGTTGTTGGTGG 660
Db |||||
601 TATCAAGTTAAGTTTGGTGTCACTACCTCTAAATGTTAATACTATGTTGTTGGTGG 660
|||

QY 661 GGTATTATTACCAATAGGTTGTAGGTGATAGGAGTAGAGAGTAGAATACTACCATCT 720
Db |||||
661 GGTATTATTACCAATAGGTTGTAGGTGATAGGAGTAGAGAGTAGAATACTACCATCT 720
|||

QY 721 ACTGCATTATCTGACGTTCTTAGAACTACTTTCAGCTTCTGCTACTTTAAATACTGATTAT 780
Db |||||
721 ACTGCATTATCTGACGTTCTTAGAACTACTTTCAGCTTCTGCTACTTTAAATACTGATTAT 780
|||

QY 781 TTTGGTTGGGAGATTGGATTAGATTGGCGTA 813
Db |||||
781 TTTGGTTGGGAGATTGGATTAGATTGGCGTA 813
|||

RESULT 6
US-10-731-554-45
; Sequence 45, Application US/10731554
; Publication No. US20040247616A1
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: McBride, Jere W.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein
; FILE REFERENCE: D6152CIP2
; CURRENT APPLICATION NUMBER: US/10/731,554
; CURRENT FILING DATE: 2003-12-09
; PRIOR APPLICATION NUMBER: US/09/611,007
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 09/660,587
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 45
; LENGTH: 813
; TYPE: DNA
; ORGANISM: Ehrlichia canis
; FEATURE:
; OTHER INFORMATION: nucleic acid sequence of E. canis p28-9
US-10-731-554-45

Query Match 100.0%; Score 813; DB 8; Length 813;
Best Local Similarity 100.0%; Pred. No. 9.2e-181;
Matches 813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAATTTACAAAAGATTGTTGAGGTGTTAGCTGTAGCTGAGTACATTTGTTTTTCTTATCT 60
Db 1 ATGAATTTACAAAAGATTGTTGAGGTGTTAGCTGTAGCTGAGTACATTTGTTTTTCTTATCT 60
|||

QY 61 GATGGTGTCTTTTCTGATGCAAAATTTTCTGAAGGAGGAGGAGGACTTTATAGTAGT 120
Db 61 GATGGTGTCTTTTCTGATGCAAAATTTTCTGAAGGAGGAGGAGGACTTTATAGTAGT 120
|||

QY 121 CAGTATAAAGTTGGTATCCCAATTTTATAGTAATTTTTCAGCTGAAGAAACAATTCCTGT 180
Db 121 CAGTATAAAGTTGGTATCCCAATTTTATAGTAATTTTTCAGCTGAAGAAACAATTCCTGT 180
|||

QY 181 ATTACAAAAGATTGTTGCGTTAGTCTTGATTAAGTCTTGAGATAAATACTCACAGCAAT 240
Db 181 ATTACAAAAGATTGTTGCGTTAGTCTTGATTAAGTCTTGAGATAAATACTCACAGCAAT 240
|||

QY 241 TTTACACGATCATATGACCCCTACTTATGCAAGCACTTTTGCAGGGTTAGTGTATCAT 300
Db 241 TTTACACGATCATATGACCCCTACTTATGCAAGCACTTTTGCAGGGTTAGTGTATCAT 300
|||

QY 301 GGATATTATGTTAAATGACTTTTATAGGTTAGGTTAGGTTAGGTTAGGTTAGGTTAGG 360
Db 301 GGATATTATGTTAAATGACTTTTATAGGTTAGGTTAGGTTAGGTTAGGTTAGGTTAGG 360
|||
```





```
Best Local Similarity 53.8%; Pred. No. 9.4e-28;
Matches 432; Conservative 0; Mismatches 353; Indels 18; Gaps 4;

QY 22 GTAGGTGTTACGCTGAGTACATTTGTTTTTCTTATCTGATGGTCTTTTCTGATGCA 81
   |||||
Db 34 GTACTTGCATCTCTATTATCTTATCTATTGAATCCTTTTCAGCTATAAATCATAAT 93

QY 82 AATTTTTCTGAAGGAGGAGACCTTATATAGGTAGTCAGTATAAAGTTGGTATCCC 141
   |||||
Db 94 CATACAGGAATAAACAATGTTGATATATATACAGGCGAGTATAGACAGGAGTATCC 153

QY 142 AATTTTAGTAATTTTTCAGCTGAAGAAACAATTTCTCGTATTACAAAAGATTTTGGG 201
   |||||
Db 154 CATTTTAGCAATTTCTCAGTAAAGAACTAATGTTGATACAACTACACTAGTAGATAT 213

QY 202 TTAGGCTTGTATAGTCT---GAGATAAATCTACAGCAATTTTACAGCATATAGAC 258
   |||||
Db 214 AAAAAAAGTGGCTCTTCTATCGATCCTAACACTTATTCAAACCTTTCAAGGTCATATACT 273

QY 259 CCTACTTTATGCAAGCAGTTTTCAGGCTTTAGTGGTATCATTTGGATATTGTTAATGAC 318
   |||||
Db 274 GTTACATTTTCAGATAAATGCTGCTAGTAAAGAACTAATGCTGCTTACTACTGTT 333

QY 319 TTT---AGGCTAGAAATTTGAAGCTTCTTATGAGAAATTTTGAACCTGAAAGACAAATGGTAC 375
   |||||
Db 334 AGCTAAGACTTGAACCTTGAAGTTCTTACGAAAAATTTTGATGTCAAAGATCCTAAAGAC 393

QY 376 CCTGAGAAATAGCAAAAGCTACAAATTTTGTCTGCGAAATGCTGCAAAATAGTAT 435
   |||||
Db 394 TACTCAGCAAAAGATGCTTTTAGGTTTTTGTCTAGCACGTAATACGCTACTACTGTT 453

QY 436 AAT-----AGTTTATAGTACTAGAGAAATACGGCGTTGTTGCAAGTCTCTTAAT 486
   |||||
Db 454 CCTGATGCTCAAAAATATACAGTTAATGAAGAAATATGCTTATCTGTTGCATCAATCATG 513

QY 487 GTAAATGTTTGTATGATATTGCTAGTGGTATGTTCTCTTTAGCACCTTATATGTTGCT 546
   |||||
Db 514 ATCAATGTTGTTATGATCTATCTTTTAAATTTAGTCTGATCACCTTATATATGTCGA 573

QY 547 GGTGTTGTCAGATATATAAAGTTTATAGGTATATGCTGCAAAATGCTACAAATAGTAT 606
   |||||
Db 574 GGTATTTGGTGAAGATTTCAATTTGATTTTGTGATCTTTGACACATTTAAACCTTGCTTATCAA 633

QY 607 GTTAAGTTTGGTCAACTACCCTTAATGTTAATGTTAATGTTGTTGCTGGGCTTAT 666
   |||||
Db 634 GGAAACTAGGTATTAGTTATTACTCTTTCTTAAGATTAAATGTTGCTGGGCTATC 693

QY 667 TACCATAAGGTTGTAGGTGATAGGATAGAGAGTAGAATAATAGCTTACCATCCTACTGCA 726
   |||||
Db 694 TATCATAGTTATAGGGAATAAATTTAAATAATTT---AAATGTTAAACCATGTTGTTACA 750

QY 727 TTATCTGACGTTCTTACAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 786
   |||||
Db 751 CTTGATGAATTTCTTAAAGCAACTTCTGCAAGTACACTTAAATGTTGCTTATTTTGGT 810

QY 787 TGGGAGATTGGATTTAGATTTCG 809
   |||||
Db 811 GGTGAAGCTGGAGTAAAGTTTAC 833
```

RESULT 9

```
US-10-901-714-23
; Sequence 23, Application US/10901714
; Publication No. US2004026533A1
; GENERAL INFORMATION:
; APPLICANT: RIKIHISA, YASUKO
; APPLICANT: OHASHI, NORIO
; TITLE OF INVENTION: OUTER MEMBRANE PROTEIN OF EHRlichia CANIS AND EHRlichia
; TITLE OF INVENTION: CHAFFEENSIS
; FILE REFERENCE: 22727-04109
; CURRENT APPLICATION NUMBER: US/10/901,714
; CURRENT FILING DATE: 2004-07-29
; PRIOR APPLICATION NUMBER: 09/314,701
```

```
; PRIOR FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: 60/100,843
; PRIOR FILING DATE: 1998-09-18
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 23
; LENGTH: 840
; TYPE: DNA
; ORGANISM: Ehrlichia chaffeensis
; US-10-901-714-23

Query Match 20.0%; Score 162.2; DB 8; Length 840;
Best Local Similarity 53.8%; Pred. No. 9.4e-28;
Matches 432; Conservative 0; Mismatches 353; Indels 18; Gaps 4;

QY 22 GTAGGTGTTACGCTGAGTACATTTGTTTTTCTTATCTGATGGTCTTTTCTGATGCA 81
   |||||
Db 34 GTACTTGCATCTCTATTATCTTATCTATTGAATCCTTTTCAGCTATAAATCATAAT 93

QY 82 AATTTTTCTGAAGGAGGAGGAGCTTTATATAGGTAGTCAGTATAAAGTTGGTATCCC 141
   |||||
Db 94 CATACAGGAATAAACAATGTTGATATATATTAACAGGCGAGTATAGACAGGAGTATCC 153

QY 142 AATTTTAGTAATTTTTCAGCTGAAGAAACAATTTCTCGTATTACAAAAGATTTTGGG 201
   |||||
Db 154 CATTTTAGCAATTTCTCAGTAAAGAACTAATGTTGATACAACTACACTAGTAGATAT 213

QY 202 TTAGGCTTGTATAGTCT---GAGATAAATCTACAGCAATTTTACAGCATATAGAC 258
   |||||
Db 214 AAAAAAAGTGGCTCTTCTATCGATCCTAACACTTATTCAAACCTTTCAAGGTCATATACT 273

QY 259 CCTACTTTATGCAAGCAGTTTTCAGGCTTTAGTGGTATCATTTGGATATTGTTAATGAC 318
   |||||
Db 274 GTTACATTTTCAGATAAATGCTGCTAGTAAAGAACTAATGCTGCTTACTACTGTT 333

QY 319 TTT---AGGCTAGAAATTTGAAGCTTCTTATGAGAAATTTTGAACCTGAAAGACAAATGGTAC 375
   |||||
Db 334 AGCTAAGACTTGAACCTTGAAGTTCTTACGAAAAATTTTGATGTCAAAGATCCTAAAGAC 393

QY 376 CCTGAGAAATAGCAAAAGCTACAAATTTTGTCTGCGAAATGCTGCAAAATAGTAT 435
   |||||
Db 394 TACTCAGCAAAAGATGCTTTTAGGTTTTTGTCTAGCACGTAATACGCTACTACTGTT 453

QY 436 AAT-----AGTTTATAGTACTAGAGAAATACGGCGTTGTTGCAAGTCTCTTAAT 486
   |||||
Db 454 CCTGATGCTCAAAAATATACAGTTAATGAAGAAATATGCTTATCTGTTGCATCAATCATG 513

QY 487 GTAAATGTTTGTATGATATTGCTAGTGGTATGTTCTCTTTAGCACCTTATATGTTGCT 546
   |||||
Db 514 ATCAATGTTGTTATGATCTATCTTTTAAATTTAGTCTGATCACCTTATATATGTCGA 573

QY 547 GGTGTTGTCAGATATATAAAGTTTATAGGTATATGCTGCAAAATGCTACAAATAGTAT 606
   |||||
Db 574 GGTATTTGGTGAAGATTTCAATTTGATTTTGTGATCTTTGACACATTTAAACCTTGCTTATCAA 633

QY 607 GTTAAGTTTGGTCAACTACCCTTAATGTTAATGTTAATGTTGTTGCTGGGCTTAT 666
   |||||
Db 634 GGAAACTAGGTATTAGTTATTACTCTTTCTTAAGATTAAATGTTGCTGGGCTATC 693

QY 667 TACCATAAGGTTGTAGGTGATAGGATAGAGAGTAGAATAATAGCTTACCATCCTACTGCA 726
   |||||
Db 694 TATCATAGTTATAGGGAATAAATTTAAATAATTT---AAATGTTAAACCATGTTGTTACA 750

QY 727 TTATCTGACGTTCTTACAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 786
   |||||
Db 751 CTTGATGAATTTCTTAAAGCAACTTCTGCAAGTACACTTAAATGTTGCTTATTTTGGT 810

QY 787 TGGGAGATTGGATTTAGATTTCG 809
   |||||
Db 811 GGTGAAGCTGGAGTAAAGTTTAC 833
```

RESULT 10

```

US-10-901-774-23
; Sequence 23, Application US/10901774
; Publication No. US20040265334A1
; GENERAL INFORMATION:
; APPLICANT: OHASHI, NORIO
; TITLE OF INVENTION: OUTER MEMBRANE PROTEIN OF EHRlichia CANIS AND EHRlichia
; TITLE OF INVENTION: CHAFFEENSIS
; FILE REFERENCE: 22727-04109
; CURRENT APPLICATION NUMBER: US/10/901.774
; PRIOR FILING DATE: 2004-07-29
; PRIOR APPLICATION NUMBER: 09/314,701
; PRIOR FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: 60/100,843
; PRIOR FILING DATE: 1998-09-18
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 23
; LENGTH: 840
; TYPE: DNA
; ORGANISM: Ehrlichia chaffeensis
US-10-901-774-23

Query Match      20.0%; Score 162.2; DB 8; Length 840;
Best Local Similarity 53.8%; Pred. No. 9.4e-28;
Matches 432; Conservative 0; Mismatches 353; Indels 18; Gaps 4;

Qy 22 GTAGGTGTACGCTGAGTACATTTGTTTTTCTTATCTGATGCTGCTTTTCTGTGATGCA 81
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 34 GTACTTGCATCTTATTCATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTAT 93
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 82 AATTTTCTGAAGGAGGAGGAGCTTATATATAGGTAGTGTAGTATAGTATAGTATAGTATAGT 141
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 94 CATACAGGAATAACACTAGTGTGTATATATATACAGGCGAGTATAGACAGGAGTATCC 153
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 142 AATTTTGTATATTTTTCAGCTGAGAACAACTTCTGGTATATACAAAAGATTTTTCGCG 201
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 154 CATTTTGTAGCAATTTTCTCAGTAAAGAACTAATGTTGATACAACTAGTAGGATAT 213
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 202 TTAGTCTTGTATAGTCT---GAGATAAATACTCACAGCAATTTTACACGATCATATGAC 258
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 214 AAAAAAGTGGCTTCTATCGATCCCTACACTTATTTCAACTTTCAAGGTCATATACT 273
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 259 CCTACTTATGCAAGCAGTTTTCAGGGTTTATGAGTATCATTTGATGATATATGTTAATGAC 318
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 274 GTTACATTTCAAGATAATGCTGTAGTTTCAGTGGGCAATTTGGATATTTCTTACCCGAA 333
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 319 TTT---AGGTAGATTTGAAGTTCTTATGAGATTTTGAACCTGGAAGCAATGGTAC 375
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 334 AGTCTAAGACTTGAACCTTGAAGTTCTTTACGAAAATTTGATGTCAAAGATCTTAAAGAC 393
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 376 CCTGAGATAGCCAAAGCTACAAATTTTTCGTTTCTCGAAATGCTTACAAATAGTGAT 435
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 394 TACTCAGCAAAAGATGCTTTTAGGTTTTCGTTCTAGCAAGTATACGTTCTACTACTGTT 453
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 436 AAT-----AAGTTTATAGTACTAGAGAAATACCGCGTTGTGTGCAAGCTCTTAAAT 486
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 454 CCTGATGCTCAAAAATATACAGTTATGAAGAATAATGGCTTATCTGTTGCATCAATCATG 513
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 487 GTAATGTTGTTATGATATGCTAGTGGTAGTATTCCTTTAGCACCTTATATGTTGCT 546
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 514 ATCAATGGTGTATGATCTATCTTTTAAATTAATTTAGTCGTATCACCTTATATATGTC 573
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 547 GGTGTTGGTGCAGATATATAAGTTTATGATATATCATTTGCCATGCTTTCTTATCAA 606
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 574 GGTATGGTGAAGATTCATGAATTTTGTATGATTTTGCATTTAACTTCTTATCAA 633
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 607 GTTAAAGTTTGGTGTCAACTACCTCTAAATGTTAATACTATGTTGTTGGTGGGGTTAT 666
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 634 GGAAACTAGGTATTAGTTATTTACTTCTTCTAAGATTAATGATTTGCTGGTGGGTAC 693
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 667 TACCATAAGTTTGTAGGTATAGGCAATGAGAGATAGAAATAGCTTACCATCTACTGCA 726
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 694 TATCATAGAGTTATAGGGAATAAATTTTAAAAATTT---AAATGTTAAACCATGTTGTTACA 750
Qy 727 TTATCTGACGTTCTTAGAACTACTTTCAGCTTCTGCTACTTTTAAATCTGATTTATTTGGT 786
Db 751 CTTGATGAATTTTCTTAAGCAAACTTCTGCGAGTAGTACACACTTAATGTTGCTTATTTGGT 810
Qy 787 TGGGAGATTGGATTTTAGATTTCG 809
Db 811 GGTGAAGCTGGAGTAAAGTTTAC 833

RESULT 11
US-10-059-964-25
; Sequence 25, Application US/10059964
; Publication No. US20020120115A1
; GENERAL INFORMATION:
; APPLICANT: RIKIHISA, YASUKO
; APPLICANT: OHASHI, NO. US20020120115A1
; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
; TITLE OF INVENTION: Chaffeensis
; FILE REFERENCE: 22727/04021
; CURRENT APPLICATION NUMBER: US/10/059,964
; CURRENT FILING DATE: 2002-01-28
; EARLIER APPLICATION NUMBER: 09/314,701
; EARLIER FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Ehrlichia chaffeensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(852)
US-10-059-964-25

Query Match      18.5%; Score 150.8; DB 5; Length 852;
Best Local Similarity 54.8%; Pred. No. 4.5e-25;
Matches 402; Conservative 0; Mismatches 307; Indels 27; Gaps 4;

Qy 96 GAGGAGAGGAGCTTTATATAGGTAGTACAGTATAAAGTTGGTATTTCCCAATTTTAGTAATTT 155
Db 111 GTCACTGGGCTATATGTCAGTCGACAAATATAAACCCTACTGTTTCTCACCTTTAGTAATTT 170
Qy 156 TTCACTGGAAGAAACAACTTCTGCTGATATTAACAAAAGATTTTTCGCTGTTAGTCTTGATA 215
Db 171 TTCACTTAAAGAACTTATATACGACAAAGAGTTATTAGGACTAGCAAAAGATATATA 230
Qy 216 GTCT-----GAGATAAATACTCACAGCAATTTTACAGGATCATATGACCTTACTTATGC 269
Db 231 GTCTATTACAGATATAACAAATAAATAAATTCACATTTCTTTATTAACAAATTTTCA 290
Qy 270 AAGCAGTTTTCAGGGGTTTAGTGTATCATTTGGATATTTATTTAATGACTTT---AGGGT 326
Db 291 AGATAATGCTGTTAGCTTTCAGTCAGCTGTTGGATATATTTCCCAAGACAGTCCCAAGGT 350
Qy 327 AGAATTTGAAGGTTCTTATGAGAAATTTTGAACCTGAAAGCAATGTTGACCTTACCAATAG 386
Db 351 TGAGGTAGAATGCTTATGAGAAATTTGACGTTTAAAAATCTCGTGAATTTACGTTAGTAAG 410
Qy 387 CCAAGCTACAAATTTTTCGTTTCTCGAAATGCTTACAAATAGTAT----- 435
Db 411 TGAAGCCTTCAGGTATATGCTTTTAGCAAGAGAAATGATAATTTCTTCAAAATATCTCTGA 470
Qy 436 ---AATAAGTTTATAGTACTAGAGATAACGGGTTGTTGACAAGTCTCTTAAATGTAAA 491
Db 471 AACAAATAGTATGTTGTTTAAAGAACATGCTTATCTGTCGATCCATTAATAACAA 530
Qy 492 TGTGTTGATGATATGCTAGTGTAGTATTCCTTTAGCACCTTATATGTTGCTGCTGTT 551
Db 531 TGGCTGTTATGATTTTCTTTTAAACAATTTTAAAGATATCACCTTACATATGCGTAGGTT 590
Qy 552 TGTGCGAGATTATATAAAGTTTTTAGGTATATCATTCGCTTAAAGTTTCTTATCAAGTTAA 611

```



Qy 327 AGAATTTGAAGGTTCTTATGAGATTTTGAACCTGGAAGCAATGGTACCTCGATAG 386  
 Db 351 TGAGGTAGAATGGTCTTATGAGAAATTTGACGTTTAAATAATCCCTGGTAAATACGTAGTAAG 410  
 Qy 387 CCAAGCTACAAATTTTTCCTTGTCTCGAAATGCTACAAATAGTGAT 435  
 Db 411 TGAAGCCTCAGGTATATGCTTTAGCAAGGNAATGATATCTTCAAAAATATCCTGA 470  
 Qy 436 ----AATAAGTTTATAGTACTAGAGAATAACGGCGTTTGTGACAAGTCTCTTAATGTA 491  
 Db 471 AACAAATAAGTATGTTGTTATAAAGAACAAATGGCTTATCTGTCGATCCATTAATCAA 530  
 Qy 492 TGTGTTATGATATTCCTAGTGTAGTATTCCTTTAGCACCTTATATGTTGCTGCTGTT 551  
 Db 531 TGGCTGTTATGATTTTCTTTAAACAATTTAAAGATATCACTTACATATGCGTAGGGTT 590  
 Qy 552 TGTGTCAGATTTATAAAGTTTATAGGTATATCAATTCCTAAGTTTCTTTATCAAGTTAA 611  
 Db 591 TGTGCGGACATATAGAAATTTTATAGTGTGTAAGTTTAAATTTGCTTATCAAGTTAA 650  
 Qy 612 GTTTGGTGTCACTACCTCTAAATGTTAACTATGTTGTTGGTGGGGTTTATTACCA 671  
 Db 651 GGTAGGTATCAGTTATCAATTTATCTCTAATATGATTTATTTGCTGACGATATTACCA 710  
 Qy 672 TAAAGTTGTAGGTATGAGGATGAGAGTAGAATAAGTATACCTACTGCAATTATC 731  
 Db 711 TAAAGTTCATAGGAATAAATTTAAACAATTTAAATGTTCAACACGTTTGTAGTCTTAA 770  
 Qy 732 TGAGTTCCTAGAACTACTCTCAGCTCTGCTGCTTAAATTAATGTTTGGTGGGGA 791  
 Db 771 TCA---TCCTAAGTCTACTTTGTCAGTAGTACTCTTAAAGTTGAGTATTTGCGTAGTGA 827  
 Qy 792 GATTGGATTTAGATTT 807  
 Db 828 ATTGGGTAAATTT 843

RESULT 14

US-10-901-774-25  
 ; Sequence 25, Application US/10901774  
 ; Publication No. US20040265334A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: RIKIHISA, YASUKO  
 ; APPLICANT: OHASHI, NORIO  
 ; TITLE OF INVENTION: OUTER MEMBRANE PROTEIN OF EHRlichia CANIS AND EHRlichia  
 ; TITLE OF INVENTION: CHAFFEENSIS  
 ; FILE REFERENCE: 2727-04109  
 ; CURRENT APPLICATION NUMBER: US/10/901,774  
 ; CURRENT FILING DATE: 2004-07-29  
 ; PRIOR APPLICATION NUMBER: 09/314,701  
 ; PRIOR FILING DATE: 1999-05-19  
 ; PRIOR APPLICATION NUMBER: 60/100,843  
 ; PRIOR FILING DATE: 1998-09-18  
 ; NUMBER OF SEQ ID NOS: 69  
 ; SOFTWARE: PatentIn Ver. 3.2  
 ; SEQ ID NO 25  
 ; LENGTH: 852  
 ; TYPE: DNA  
 ; ORGANISM: Ehrlichia chaffeensis  
 US-10-901-774-25

Query Match 18.5%; Score 150.8; DB 8; Length 852;  
 Best Local Similarity 54.6%; Pred. No. 4.5e-25;  
 Matches 402; Conservative 0; Mismatches 307; Indels 27; Gaps 4;

Qy 96 GAGGAGAGACTTTATATAGTAGTACGATATAAAGTTGGTATCCCAATTTTAGTAATTT 155  
 Db 111 GTGCACCTGGCTATATGTCAGTGGCAATATAAACTACTGTTCTCTCACTTTAGTAATTT 170  
 Qy 156 TTCAGTCTAGAAACAATTCCTGGTATTAACAAAAGATTTTGGCTTAGTCTTGATAA 215  
 Db 171 TTCACCTAAGAACTTTATGACTGACACTAAAGAGTTATTAGGACTAGCAAAAGATTTAA 230

Qy 216 GTCT-----GAGATAAATACTCAGACAAATTTTACACGATCATATGACCTACTTATGC 269  
 Db 231 GTCTATTACAGATATTAACAACAAATTAATAATTCACATTCCTTATAACACAAAATTTCA 290  
 Qy 270 AAGCAGTTTTCAGGGTTTGTAGTATCAATGGATATTAATGTTAAATGACTTTT---AGGGT 326  
 Db 291 AGATAATGCTGTTAGCTTCAGTGCAGCTGTTGGATATATTTCCCAAGACAGTCCCAAGGGT 350  
 Qy 327 AGAATTTGAAGTTCTTATGAGAAATTTTGAACCTGAAAGACAAATGCTGACCTGAGAAATAG 386  
 Db 351 TGAGGTAGAAATGCTCTTATGAAGAATTTGACGTTAAATAATCCTGGTAAATTTACGTAGTAAG 410  
 Qy 387 CCAAGCTACAAATTTTTCCTGCTCGAAATGCTACAAATAGTGAT 435  
 Db 411 TGAAGCCTTCAGGTATATGCTTTTAGCAAGAGGAATGATATCTTCAAAAATATCCTGA 470  
 Qy 436 ----AATAAGTTTATAGTACTAGAGAATAACGGCGTTTGTGACAAGTCTCTTAAATGTA 491  
 Db 471 AACAAATAAGTATGTTGTTATAAAGAACAAATGCTTATCTGTCGATCCATTAATCAA 530  
 Qy 492 TGTGTTATGATATTTGCTAGTGTAGTATTCCTTTAGCACCTTATATGTTGCTGCTGTT 551  
 Db 531 TGGCTGTTATGATTTTCTTTAAACAATTTAAAGATATCACCTTACATATGCGTAGGGTT 590  
 Qy 552 TGTGTCAGATTTATAAAGTTTTCCTAGGTATATCAATTCCTTAAGTTTCTTTATCAAGTTAA 611  
 Db 591 TGTGCGGACATTTATAGAAATTTTATAGTGTGCTGCTGTAAGTTTAAATTTGCTTATCAAGTTAA 650  
 Qy 612 GTTTGGTGTCACTACCTCTTAAATGTTAACTATGTTGTTGGTGGGGTTTATTACCA 671  
 Db 651 GGTAGGTATCAGTTATCCATTTATCTCTAATATGATTTATTTGCTGACGATATTACCA 710  
 Qy 672 TAAAGTTGTAGGTATGAGGATGAGAGTAGAATAAGTATACCTACTGCAATTATC 731  
 Db 711 TAAAGTTCATAGGAATAAATTTAAACAATTTAAATGTTCAACACGTTGTTAGTCTTAA 770  
 Qy 732 TGAGTTCCTAGAACTACTCTCAGCTTCGCTACTCTTAAATTAATGATTATTTGGTGGGGA 791  
 Db 771 TCA---TCCTAAGTCTACTTTTGCAGTAGTACTCTTAAATGTTGAGTATTTGCGTAGTGA 827  
 Qy 792 GATTGGATTTAGATTT 807  
 Db 828 ATTGGGTAAATTT 843

RESULT 15

US-09-811-007-41  
 ; Sequence 41, Application US/09811007  
 ; Publication No. US20030185849A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Walker, David H.  
 ; APPLICANT: McBride, Jere W.  
 ; APPLICANT: Yu, Xue-Jie  
 ; TITLE OF INVENTION: Homologous 28-kilodalton Immunodominant Protein  
 ; TITLE OF INVENTION: Genes of Ehrlichia canis and Uses Thereof  
 ; FILE REFERENCE: D6152CIP2  
 ; CURRENT APPLICATION NUMBER: US/09/811,007  
 ; CURRENT FILING DATE: 2001-03-16  
 ; PRIOR APPLICATION NUMBER: 09/660,587  
 ; PRIOR FILING DATE: 2000-09-12  
 ; NUMBER OF SEQ ID NOS: 46  
 ; SEQ ID NO 41  
 ; LENGTH: 840  
 ; TYPE: DNA  
 ; ORGANISM: Ehrlichia canis  
 ; FEATURE:  
 ; OTHER INFORMATION: nucleic acid sequence of E. canis p28-2  
 US-09-811-007-41

Query Match 18.1%; Score 147.2; DB 3; Length 840;  
 Best Local Similarity 53.2%; Pred. No. 3.1e-24;  
 Matches 445; Conservative 0; Mismatches 358; Indels 33; Gaps 5;

Qy 1 ATGAATTACAAAGATTGTTGTAGGTGTTACGCTGAGTACATTGTTTTTTCTTATCT 60  
 |||||  
 Db 1 ATGAATTATAAGAAATCTAGTAAGAAGCGCTAATCTCATTAATGTCAATCTTACCA 60  
 |||||  
 Qy 61 GATGGTGCCTTTTCTGA-----TGCAATTTTCTGAAGGAGGAGAGACTTTAT 111  
 |||||  
 Db 61 TATCAGTCTTTTGCAGATCCTGTAGGTTCAAGAACTAATGATAACAAGAAGGCTTCTAC 120  
 |||||  
 Qy 112 ATAGGTAGTCAGTATAAGTTGGTATTCCTCAATTTTAGTAATTTTTCAGCTGAGAAACA 171  
 |||||  
 Db 121 ATTAGTGAAGTACAAATCCAGTATATACACTTTTAGAAAATTTCTCTGCTGAAGAACT 180  
 |||||  
 Qy 172 ATTCCTGGTATTACAAAAAGATTTTTCGTTAGGCTTGG-----ATAAGTCTGAG 222  
 |||||  
 Db 181 CCTATTAAATGGACAAATCTCTCACTAAAAAGTTTTCGGACTAAAGAAAGATGGTGAT 240  
 |||||  
 Qy 223 ATAAATCTCACAGCAATTTTACAGATCATATGACCCCTACTTAT-----GCAAGCAGT 276  
 |||||  
 Db 241 ATAACAAAAAAGACGATTTTACAAGAGTAGCTCCAGGCATTGATTTTCAAAATAACTTA 300  
 |||||  
 Qy 277 TTTGCGAGGTTTAGTGATCATTTGGATATTATGTTAATGACTTTTAGGGTAGAATTTGAA 336  
 |||||  
 Db 301 ATATCAGGATTTTCAGGAAGTATTTGGTTACTCTATGGACGGACCAAGAATAGAACTTGAA 360  
 |||||  
 Qy 337 GGTTCCTTATGAGAATTTTGAACCTGAAAGACAATGGTACC-----CTGAGAAATAGCCAA 390  
 |||||  
 Db 361 GCTGCATATCAACAAATTTATCCAAAAACACCGATAACAATGATCTGATATATGGTGAA 420  
 |||||  
 Qy 391 AGTACAAATTTTGTCTGCGAAATGCTACAAATGCTACAAATAGTGATTAATAGTTTATAGTA 450  
 |||||  
 Db 421 TACTATAAACATTTTGCAATTATCTCGTAAAGATGCAATGGAAGATCAGCAATATGTAGTA 480  
 |||||  
 Qy 451 CTAGAGAAATACCGCGTTGTTGACAACTCTTTAATGTAATGTTTGTATGATATTGCT 510  
 |||||  
 Db 481 CTTAAAAATGACGGCATAACTTTTATGTCATTGATGGTTAATACTTGTCTATGACATTACA 540  
 |||||  
 Qy 511 AGTGGTAGTATTCCTTTAGCACCTTATATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 570  
 |||||  
 Db 541 GCTGAGGAGTATCTTTTCGTACCATATGCTATGCTAGGTATAGGAGCAGATCTTATCACT 600  
 |||||  
 Qy 571 TTTTGTAGTATATCATGCTCAAGTTTCTTATCAAGTTAAGTTTGGTGTCAACTACCT 630  
 |||||  
 Db 601 ATTTTAAAGACCTCAATCTAAAATTTGCTTACCAAGGAAAAATAGGTATTAGTTACCT 660  
 |||||  
 Qy 631 CTAAATGTTAATACTATGTTGTTGGTGGGGTTATTACCATAGGTTGTAGTGATAGG 690  
 |||||  
 Db 661 ATCACACCAAGTCTCTGCAATTATTTGGTGGATACTTACCATGGCGTTATTTGGTAATAA 720  
 |||||  
 Qy 691 CATGAGAGAGTAGAATAGCTTACCATCTCTGCTATCTGACGTTCTAGAACTACT 750  
 |||||  
 Db 721 TTTGAGAGATACCTGTA---ATAACTCTCTGTAGTATTAATGATGCTCTCAACACCA 777  
 |||||  
 Qy 751 TCAGCTTCTGCTATTTTAAATCTGATATTTTGGTTGGGAGATTGGATTAGATT 806  
 |||||  
 Db 778 TCTGCTTCAGTAACTCTTGACGTTGGATCTTTGGGGAGAAAATTGGAAATGAGGTT 833  
 |||||

Search completed: December 28, 2005, 15:20:21  
 Job time : 872 secs